

FORM PTO-1390 (REV. 10-2000)		U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE	ATTORNEY'S DOCKET NUMBER 20377YP
TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371			U.S. APPLICATION NO. (If known, see 37 CFR 1.5) 09/890836
INTERNATIONAL APPLICATION NO. PCT/US00/02405	INTERNATIONAL FILING DATE 31 JANUARY 2000	PRIORITY DATE CLAIMED 04 FEBRUARY 1999	
TITLE OF INVENTION IMPROVED HELPER DEPENDENT VECTOR SYSTEM FOR GENE THERAPY			
APPLICANT(S) FOR DO/EO/US ANDREW BETT, VOLKER SANDIG, AND RIMA YOUIL.			
Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:			
<ol style="list-style-type: none"> <input checked="" type="checkbox"/> This is a FIRST submission of items concerning a filing under 35 U.S.C. 371. <input type="checkbox"/> This is a SECOND or SUBSEQUENT submission of items concerning a filing under 35 U.S.C. 371. <input type="checkbox"/> This is an express request to begin national examination procedures [35 U.S.C. 371(f)] at any time rather than delay examination until the expiration of the applicable time limit set in 35 U.S.C. 371(b) and PCT Articles 22 and 39(1). <input type="checkbox"/> A proper Demand for International Preliminary Examination was made and the US was elected by the expiration of the 19th month from the earliest claimed priority date (PCT Article 31). <input checked="" type="checkbox"/> A copy of the International Application as filed [35 U.S.C. 371(c)(2)]. <ol style="list-style-type: none"> <input type="checkbox"/> is attached hereto (required only if not communicated by the International Bureau). <input type="checkbox"/> has been communicated by the International Bureau. <input checked="" type="checkbox"/> is not required, as the application was filed in the United States Receiving Office (RO/US). <input type="checkbox"/> An English language translation of the International Application as filed [35 U.S.C. 371(c)(2)]. <input type="checkbox"/> Amendments to the claims of the International Application under PCT Article 19 [35 U.S.C. 371(c)(3)]. <ol style="list-style-type: none"> <input type="checkbox"/> are attached hereto (required only if not communicated by the International Bureau). <input type="checkbox"/> have been communicated by the International Bureau. <input type="checkbox"/> have not been made; however, the time limit for making such amendments has NOT expired. <input type="checkbox"/> have not been made and will not be made. <input type="checkbox"/> An English language translation of the amendments to the claims under PCT Article 19 [35 U.S.C. 371(c)(3)]. <input checked="" type="checkbox"/> An oath or declaration of the inventor(s) [35 U.S.C. 371(c)(4)]. <input type="checkbox"/> An English language translation of the annexes to the International Preliminary Examination Report under PCT Article 36 [35 U.S.C. 371(c)(5)]. 			
Items 11 to 16 below concern other document(s) or information included:			
<ol style="list-style-type: none"> <input type="checkbox"/> An Information Disclosure Statement under 37 CFR 1.97 and 1.98. <input type="checkbox"/> An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included. <input checked="" type="checkbox"/> A FIRST preliminary amendment. <input type="checkbox"/> A SECOND or SUBSEQUENT preliminary amendment. <input type="checkbox"/> A substitute specification. <input type="checkbox"/> A change of power of attorney and/or address letter. <input type="checkbox"/> Other items or information: 			

EXPRESS MAIL CERTIFICATE
DATE OF DEPOSIT Aug 3, 2001
EXPRESS MAIL NO. EL052390686505
I HEREBY CERTIFY THAT THIS CORRESPONDENCE IS
BEING DEPOSITED WITH THE UNITED STATES POSTAL
SERVICE AS EXPRESS MAIL "POST OFFICE TO
ADDRESSEE" BEFORE 5 P.M. ON THE ABOVE DATE IN
AN ENVELOPE ADDRESSED TO ASSISTANT COMMISSIONER
FOR PATENTS, WASHINGTON, D.C. 20231
MAILED BY Robert R. Kutzowski
DATE August 3, 2001

U.S. APPLICATION NO (If known, see 37 CFR 1.5) 09/890836		INTERNATIONAL APPLICATION NO. PCT/US00/02405		ATTORNEY'S DOCKET NUMBER 20377YP	
--	--	---	--	-------------------------------------	--

17. <input checked="" type="checkbox"/> The following fees are submitted:				CALCULATIONS	PTO USE ONLY
BASIC NATIONAL FEE [37 CFR 1.492(a)(1)-(5)]: Neither international preliminary examination fee (37 CFR 1.482) nor international search fee [37 CFR 1.445(a)(2)] paid to USPTO and International Search Report not prepared by the EPO or JPO..... \$1,000.00 International preliminary examination fee (37 CFR 1.482) not paid to USPTO but International Search Report prepared by the EPO or JPO..... \$860.00 International preliminary examination fee (37 CFR 1.482) not paid to USPTO but international search fee [37 CFR 1.445(a)(2)] paid to USPTO. \$710.00 International preliminary examination fee paid to USPTO (37 CFR 1.482) but all claims did not satisfy provisions of PCT Article 33(1)-(4)..... \$690.00 International preliminary examination fee paid to USPTO (37 CFR 1.482) and all claims satisfied provisions of PCT Article 33(1)-(4) \$100.00 ENTER APPROPRIATE BASIC FEE AMOUNT =				\$100.00	
Surcharge of \$130.00 for furnishing the oath or declaration later than <input type="checkbox"/> 20 <input type="checkbox"/> 30 months from the earliest claimed priority date [37 CFR 1.492(e)].				\$0.00	
Claims	Number Filed	Number Extra	Rate		
Total Claims	41 - 20 =	21	X \$18.00	\$378.00	
Independent Claims	5 - 3 =	2	X \$80.00	\$160.00	
Multiple dependent claim(s) (if applicable)			+ \$270.00	\$0.00	
TOTAL OF ABOVE CALCULATIONS =				\$638.00	
<input type="checkbox"/> Applicant claims small entity status. See 37 CFR 1.27. The fees indicated above are reduced by 1/2.					
SUBTOTAL =				\$638.00	
Processing fee of \$130.00 for furnishing the English translation later than <input type="checkbox"/> 20 <input type="checkbox"/> 30 months from the earliest claimed priority date [37 CFR 1.492(f)].				+	\$0.00
TOTAL NATIONAL FEE =				\$638.00	
Fee for recording the enclosed assignment [37 CFR 1.21(h)]. The assignment must be accompanied by an appropriate cover sheet (37 CFR 3.28, 3.31). \$40.00 per property.				+	
TOTAL FEES ENCLOSED =				\$638.00	
				Amount to be refunded	
				charged	

a. ☐ A check in the amount of \$ _____ to cover the above fees is enclosed.

b. ☒ Please charge my Deposit Account No. 13-2755 in the amount of \$638.00 to cover the above fees.
A duplicate copy of this sheet is enclosed.

c. ☒ The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any
overpayment to the Deposit Account No. 13-2755. A duplicate copy of this sheet is enclosed.

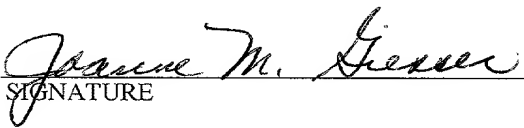
**NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive
[37 CFR 1.137(a) or (b)] must be filed and granted to restore the application to pending status.**

SEND ALL CORRESPONDENCE TO:

MERCK & CO., INC.
 Patent Department, RY60-30
 P.O. Box 2000
 126 East Lincoln Avenue
 Rahway, New Jersey 07065-0970

DATE: August 3, 2001

PHONE #: (732) 594-3046


 SIGNATURE
 JOANNE M. GIESSER
 NAME
32,838
 REGISTRATION NUMBER

PATENT

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicants: A. Bett et al.

Serial No.: TO BE ASSIGNED Case No.:20377YP.
NATIONAL FILING OF
PCT/US00/02405

Art Unit:

Filed: HEREWITH

Examiner:

For: IMPROVED HELPER-DEPENDENT
VECTOR SYSTEM FOR GENE
THERAPY

The Assistant Commissioner for Patents
Washington, D.C. 20231

PRELIMINARY AMENDMENT

Dear Sir:

Please amend the claims to read as follows:

4. (Amended) The nucleic acid of claim 1, wherein said modified packaging signal is less efficient than said wild-type packaging signal.

32. (Amended) The helper-dependent adenoviral vector of claim 28, wherein said optionally present non-coding E4 region is present.

EXPRESS MAIL CERTIFICATE
DATE OF DEPOSIT Aug 3, 2001
EXPRESS MAIL NO. EL52390686505
I HEREBY CERTIFY THAT THIS CORRESPONDENCE IS
BEING DEPOSITED WITH THE UNITED STATES POSTAL
SERVICE AS EXPRESS MAIL, "POST OFFICE TO
ADDRESSEE" BEFORE 5 P.M. ON THE ABOVE DATE IN
AN ENVELOPE ADDRESSED TO ASSISTANT COMMISSIONER
FOR PATENTS, WASHINGTON, D.C. 20231.
MAILED BY Kathy Blumenthal
DATE August 3, 2001

09/890836

41. (amended) The method of claim 36, wherein said wild-type packaging signal is from human adenovirus serotype 5.

REMARKS

This is a preliminary amendment; claims 1-41 are pending. The amendments to claims 4, 32 and 41 merely change multiply dependent claims to non-multiply dependent claims. No new matter has been added.

Attached is Appendix 1, showing changes made to each claim.

Applicants assert that all claims are in condition for allowance, and a favorable action on the merits is earnestly solicited.

Respectfully submitted,

By Joanne M. Giesser
Joanne M. Giesser
Reg. No. 32,838
Attorney for Applicants

Merck & Co., Inc.
P.O. Box 2000
Rahway, NJ 07065-0907
(732) 594-3046

Date: August 3, 2001

09890835-08004
T0E020-9E806860

8pt

TITLE OF THE INVENTION

IMPROVED HELPER DEPENDENT VECTOR SYSTEM FOR GENE THERAPY

5 FIELD OF THE INVENTION

This invention relates to adenoviral vectors which are useful for nucleic acid delivery into a cell such as in gene delivery and nucleotide vaccine applications. It also relates to host cells and methods using these vectors.

10 BACKGROUND OF THE INVENTION

The references cited herein are not admitted to be prior art to the claimed invention.

Adenoviral vectors provide a vehicle for introducing nucleic acids into a cell *in vitro* and *in vivo*. Over 100 different serotypes, of which 47 are of human origin, have been reported. (Hierholzer, *et al.*, *J. infect. Dis.* 159:804-813, 1988.)
15 Adenoviruses type 2 and 5 used to produce adenoviral vectors are well characterized.

A human adenovirus contains a 5' inverted terminal repeat (ITR); a packaging signal; region E1 made up of the early regions E1A and E1B; region E2; region E3, region E4; the late regions L1-L5; and a 3' ITR. Regions E1 and E4
20 contain regulatory proteins, region E2 encodes proteins required for replication, and the L region encodes for the structural proteins of the virus. The E3 region is dispensable for virus growth *in vitro*. (Hitt, *et al.*, *Advances in Pharmacology* 40:137-206, 1996.)

A replicating viral vector based on the adenovirus must contain
25 adenovirus *cis* elements needed for replication. Packaging also requires *cis* elements on the vector whereas the necessary proteins can be supplied in *trans*. In *trans* supplementation can be brought about by specialized cells and/or additional viruses producing the needed proteins. (Hitt, *et al.*, *Advances in Pharmacology* 40:137-206, 1996.)

30 Different types of adenoviral vectors have been developed including those lacking one or more components of the adenoviral genome. An adenoviral vector produced without components of the adenoviral genome provides advantages such as in increasing the amount of foreign nucleic acid that can be included with the adenoviral vector and removing adenoviral genes whose expression can have
35 detrimental effects. (Hitt, *et al.*, *Advances in Pharmacology* 40:137-206, 1996.)

Adenoviral vectors lacking all viral protein-coding sequences have been developed. These vectors require supplementation of viral regulatory and structural proteins supplied *in trans* for packaging and rescue. A second adenovirus carrying genes necessary for virus growth can be used to provide *in trans* the required supplementation of proteins. (Mitani, *et al.*, *Proc. Natl. Acad. Sci. USA* 92:3854-3858, 1995; Fisher, *et al.*, *Virology* 217:11-22, 1996; Kochanek, *et al.*, *Proc. Natl. Acad. Sci. USA* 93:5731-5736, 1996; Parks, *et al.*, *Proc. Natl. Acad. Sci. USA* 93:13565-13570, 1996; Parks, *et al.*, *J. Virology* 71(4):3293-3298, 1997; and Schiedner, *et al.*, *Nature Genetics* 18:180-183, 1998.)

10

SUMMARY OF THE INVENTION

The present invention features helper-dependent adenoviral vector elements, and helper adenoviral elements, that enhance the production and isolation of helper-dependent adenoviral vectors. Such elements include a modified packaging signal having low homology to, and preferably less activity than, a wild-type packaging signal, an E4 non-coding segment directly joined to the 5' ITR that confers a selective advantage, and stuffer region(s) that provide a helper-dependent adenoviral vector with a GC content of about 50% to about 60%. The modified packaging signal is preferably used in a helper virus to decrease recombination and generation of the virus. The E4 non-coding segment and the stuffer region(s) are preferably used in a helper-dependent adenoviral vector to provide the vector with a growth advantage over a helper virus.

A "helper-dependent adenoviral vector" refers to a viral vector containing the *cis* elements needed for adenovirus replication and packaging (also referred to herein as viral generation), but lacking the necessary *trans* elements for adenovirus replication and packaging. The *cis* elements needed for viral generation is an adenovirus 3' ITR, a packaging signal, and an adenovirus 5' IRT. *Trans* elements needed for viral generation are the proteins encoded by the E1, E2, and E4 and L adenovirus regions. Preferably, the helper-dependent adenoviral vector lacks nucleic acid encoding for any adenovirus proteins. A helper-dependent adenoviral vector is particularly useful as a vector for delivering genes into cells *in vivo*.

A "helper virus" refers to a virus expressing one or more proteins needed for viral generation of a helper-dependent adenovirus. Preferably, the helper virus contains the adenovirus 3' IRT, an adenovirus packaging signal, an adenovirus

5' IRT, and encodes for the proteins of the Adenovirus E2 and E4 and L regions. *Trans* elements not provided for by a helper virus may be provided for by other means such as by a cell. For example, a 293 cell can be employed to supply in *trans* needed E1 proteins in conjunction with a helper virus not expressing such proteins.

5 Helper-dependent adenoviral vectors can be produced by co-cultivating with helper viruses in cell lines, wherein the helper virus alone or in combination with the cell line provide the *trans* functions needed for adenovirus generation. During co-cultivation of a helper-dependent adenoviral vector and a helper virus the two compete for packaging, producing a mixed viral population.

10 Often, during co-cultivation the helper virus has a selection advantage, such that the mixed population of helper virus and helper-dependent adenoviral vector becomes predominantly helper virus, rather than the desired helper-dependent adenoviral vector. Even when this phenomenon does not occur, the presence of lesser, but significant, amounts of helper viruses still contaminates the helper-
15 dependent adenoviral vector preparation. Therefore, the viral populations may have to be separated physically and/or production of helper particles suppressed for the helper-dependent adenoviral vector to be used. Preferably, the percentage of helper virus in a purified helper-dependent adenoviral vector stock is below about 0.5%.

20 Thus, a first aspect of the present invention describes a nucleic acid molecule comprising an excisable low homology packaging signal cassette. The packaging cassette comprises a low homology packaging signal cassette flanked by a recombinase recognition sequence. The modified packaging signal has low homology to a wild-type adenovirus packaging signal.

25 "Flanked by a recombinase recognition sequence" indicates the presence of a recombinase recognition sequence 3' and 5' of the packaging signal allowing for excision of the flanked nucleic acid containing the packaging signal by a recombinase recognizing the recognition sequence. The recombinase recognition sequence does not need to be immediately 3' or 5' of the modified packaging signal. Preferably, the 3' and 5' recognition sequence are about 200 bp to about 2,000 bp
30 apart.

 The modified packaging signal has low homology to a wild-type adenovirus packaging signal and still allows for packaging of the helper virus during its separate production. Preferably, the modified packaging signal is less efficient than a wild-type adenovirus packaging signal.

“Low homology” refers to a maximum of about 25 bp of contiguous sequence homology (100% sequence identity) between the modified packaging signal and a wild type packaging signal present in an adenovirus, preferably human adenovirus serotype 5. Low homology is determined by aligning the modified packaging signal with wild type packaging signal to obtain a stretch of maximum contiguous homology with sequences present in both the modified and wild type packaging signals. Preferably, contiguous sequence homology is at most 23 bp.

“Less efficient” than a wild-type adenovirus packaging signal indicates that the helper virus has a lower level of packaging when it contains the modified packaging signal than it has when it contains a wild-type packaging signal present in an adenovirus, preferably from a human adenovirus group C serotype, more preferably serotype 5. Preferably, the helper virus can still be produced at yields of at least 300 PFU/cell in cells lacking a compatible recombinase. Efficiency is preferably determined using titration by end point dilution assay described in the Examples below.

Reference to “modified” is not a limitation as to how the packaging signal sequence is produced. Instead, “modified” indicates that the sequence is not a wild-type sequence.

Another aspect of the present invention describes an adenoviral helper virus for production of helper-dependent vectors comprising:

- (a) an adenovirus genome having an E1 region deletion;
- (b) an excisable packaging signal cassette replacing a wild-type packaging signal, the excisable packaging signal cassette comprising a 5' *loxP* site, a modified packaging signal and a 3' *loxP* site, wherein the modified packaging signal has low homology to and is less efficient than the wild-type packaging signal; and
- (c) an optional insertion element comprising at least about 2900 base pairs of non-adenoviral DNA inserted in the E3 region without deleting any part of the E3 region.

Another aspect of the present invention describes a helper-dependent adenovirus vector comprising the following elements (in 5' to 3' order): (a) a 5'-inverted terminal repeat sequence; (b) a packaging signal; (c) one or more heterologous gene expression cassettes; (d) an optionally present E4 non-coding segment conferring a selective advantage; and (e) a 3' ITR; wherein (d) and (e) are located in the distal 400 bp of the adenovirus genome, and wherein the only

adenoviral sequences present in the vector are said 5'-inverted terminal repeat, said 3'-inverted terminal repeat, said packaging signal, and said nonexpressed segment of the E4 region.

5 The "E4 non-coding segment" refers to the nonexpressed segment of the E4 region co-localized with the E4 promoter. The segment is adjacent to the 3' ITR and, in human adenovirus serotype 5, is about 300 bp in length.

Another aspect of the present invention describes a helper-dependent adenovirus vector having a GC content between about 50% and about 60%. The vector comprises in a 5' to 3' direction: (a) a 5' ITR; (b) a packaging signal cassette
10 directly joined to the 3' of the 5' ITR; (c) a first stuffer DNA at least about 1 kb; (d) at least one heterologous expression cassette; (e) a second stuffer DNA at least about 1 kb; (f) an optionally present non-coding E4 segment; and (g) a 3' ITR. Element (f) if present is directly joined to element (g). The virus does not encode one or more adenovirus proteins needed for adenovirus generation and is about 28 kb to about 36
15 kb. Preferably, the virus does not encode any adenovirus proteins.

"Directly joined" indicates the absence of significant intervening sequences (*e.g.*, less than about 150 bp) that interfere with the function of "joined" groups.

"Stuffer sequences" refer to nucleic acid regions that do not express
20 protein when present in a helper-dependent adenoviral vector. Preferably, such regions are mammalian non-gene regions or introns.

Another aspect of the present invention describes a cell line infected with a helper virus, a helper-dependent viral vector, or both the virus and the vector. Preferably, the cell line expresses E1 proteins and a recombinase, and is infected with
25 a helper virus that (1) contains an excisable low homology signal packaging cassette, and (2) does not express E1 proteins needed for viral generation.

Another aspect of this invention is a method of generating helper-dependent adenoviral gene vectors in a cell line expressing E1 proteins and Cre recombinase comprising:

30 a) infecting the cell line with a helper-dependent vector comprising: a 5' ITR, a packaging signal, at least one heterologous expression cassette, human genomic stuffer DNA and a 3' ITR, wherein the overall size of the helper-dependent vector is between about 28 kb and 36 kb, and wherein no functional adenoviral

coding sequences and no bacterial origin of replication or bacterial marker genes are present;

- b) infecting the cell line with a helper virus comprising: an adenovirus genome having an E1 region deletion; an excisable packaging signal cassette replacing a wild-type packaging signal, the excisable packaging signal cassette comprising a 5' *loxP* site, a modified packaging signal and a 3' *loxP* site, wherein the modified packaging signal has low homology to and is less efficient than the wild-type packaging signal; and an optional insertion element comprising at least about 2900 base pairs of non-adenoviral DNA inserted in the E3 region without deleting any part of the E3 region; and
- c) obtaining the generated helper-dependent viral vectors.

Another aspect of the present invention describes a method of generating helper-dependent adenoviral vectors. The method involves producing a cell comprising: (i) *trans* functions needed for adenovirus generation and (ii) a helper-dependent adenoviral vector comprising the necessary *cis* functions needed for adenovirus generation and at least one heterologous expression cassette, wherein the helper-dependent adenoviral vector does not encode for any adenovirus proteins, is about 28 kb to about 36 kb, and has a GC content between about 50% and about 60%. The cell is used to generate the helper-dependent vector.

Reference to a cell comprising the *trans* functions needed for adenovirus generation indicates such functions are present in the cell, but need not be part of the cellular genome. The *trans* functions can be provided, for example, by a helper virus, the cellular genome, or a combination of both. Another example of a source of *trans* functions is a plasmid.

Other features and advantages of the present invention are apparent from the additional descriptions provided herein including the different examples. The provided examples illustrate different components and methodology useful in practicing the present invention. The examples do not limit the claimed invention. Based on the present disclosure the skilled artisan can identify and employ other components and methodology useful for practicing the present invention.

BRIEF DESCRIPTION OF THE DRAWINGS

FIGURE 1 shows the shuffling of fragments of a stuffer to prevent genomic integration of the expression units.

FIGURE 2 is a gel showing restriction digests of helper-dependent adenoviral vector DNA after competitive rescues of: a) stk120gfp-E4-promoter and stk120gfp; b) stk120gfp-E4 and stk120gfp; and c) stk120gfp-E4-promoter and stk120gfp-E4.

5 FIGURE 3 is a gel showing restriction digests of helper-dependent adenoviral vectors grown with the helper AdLC8BHG and a helper of this invention.

FIGURE 4 shows the structure of different helper-dependent adenoviral vectors.

FIGURE 5 shows autoradiographs of labeled restriction fragments
10 from helper-dependent adenoviral vectors. The figure illustrates the results of competition experiments involving groups of helper-dependent adenoviral vectors from transfection to passage 6. "A" is a competition of backbones containing the shuffled fragments of cosmid HUMDXS455A (C4). "B" is a competition of backbones containing part of the HPRT (hypoxanthine guanine
15 phosphoribosyltransferase) genomic gene. "C" is a competition between 2 backbones one containing part of the HPRT genomic gene the other containing the shuffled fragments of HUMDXS455A (C4).

FIGURE 6 is a graph showing the correlation between GC content and helper virus contamination.

20 FIGURE 7 is a graph showing the balance between helper and helper-dependent adenoviral vectors at different time points after infection. The figure demonstrates the replication potential of different helper-dependent adenoviral vectors with the mouse Erythropoetin gene relative to the helper virus.

25 DETAILED DESCRIPTION OF THE INVENTION

The present invention is particularly useful for obtaining and isolating helper-dependent adenoviral vectors produced using a helper virus. Co-cultivation of a helper-dependent adenoviral vector and a helper virus produces two different particles competing for packaging and results in a mixed viral population. The
30 production of the helper virus in the mixed viral population reduces the amount of helper-dependent adenoviral vector produced, provides a viral contaminant that interferes with applications of the helper-dependent adenoviral vector, and produces a source of nucleic acid that can be recombined with helper-dependent adenoviral vectors

As described herein, production and isolation of helper-dependent adenoviral vectors can be facilitated in different ways. One way is to employ helper-dependent adenoviral vectors containing elements favoring viral generation of the helper-dependent adenoviral vector over a co-cultured helper virus. Another way is to employ helper viruses containing elements that can be employed to reduce the ability of the helper virus to grow, and /or recombine with the helper-dependent adenoviral vector.

The term "vector" employed herein is used in its broadest sense, and is meant to encompass a linear virus which can carry a heterologous gene, and or a plasmid which can contain one or more regions of a virus genome. The term encompasses a helper-dependent adenoviral vector and a helper virus.

Adenovirus components present in a helper or a helper-dependent vector may in general be provided from any adenovirus. Preferably, a human adenovirus serotype is used; more preferably, human adenovirus serotype 1-47; and more preferably, a group C serotype (*e.g.*, serotypes 1, 2, 5, and 6). Preferred group C serotypes are human adenovirus serotype 2 or 5, and more preferably serotype 5.

A helper-dependent adenoviral vector preferably contains a heterologous expression cassette. Such a cassette is made up of a heterologous gene functional coupled to a promoter and regulator elements needed for gene expression, and is particularly useful for introducing nucleic acid into a cell. The introduced nucleic acid preferably encodes for a gene or is able to regulate gene expression. Nucleic acid able to regulate gene expression include ribozymes and antisense nucleic acids.

The ability to introduce nucleic acid into a cell using the present invention has different applications such as in gene therapy, in antibody production (*e.g.*, a vaccine) and research to examine the regulation of a gene *in vivo* or *in vitro*. Different genes can be introduced into a cell such as Factor VIII, Factor VIX, CFTR, OTC, LDL, VEGF, FGF, EPO, HSV-TK, IL-2, IL-12, p53, HLA-B7, α -interferon, and cytosine deaminase.

A promoter is a DNA sequence directing the synthesis of RNA through an RNA polymerase. Suitable promoters depend upon the gene and intended use, and may include promoters such as EF1 α , CMV, chicken α -actin (Arnold, *et al.*, Nucleic Acids Res 1988 Mar 25;16(6):2411-29), and muscle creatine kinase (Johnson, *et al.*, Mol Cell Biol 1989 Aug;9(8):3393-9).

Generally, the regulatory elements that are present include a ribosome binding site, a terminator, and an optionally present operator. Another preferred element is a polyadenylation signal. Regulatory systems are available to control gene expression, such as GENE-SWITCH™ (Wang, *et al.*, *Gene Ther.* 1997 May;4(5):432-41, U.S. Patent No. 5,874,534 and International Publication WO 93/23431, each of which are hereby incorporated by reference herein) and those involving the tetracycline operator (U.S. Patent Nos. 5,464,758 and 5,650,298, both of which are hereby incorporated by reference herein).

A helper-dependent adenoviral vector is preferably from about 28 kb to about 36 kb to provide for efficient packaging. An adenovirus can accommodate up to about 105% of the wild-type genome and has a lower packaging limit of about 75% of the wild-type genome. (See, Parks, *et al.*, *J. Virology* 71(4):3293-3298, 1997, hereby incorporated by reference herein.) In a preferred embodiment, there are no bacterial plasmid-based sequences (such as an origin of replication or bacterial marker genes) present in the helper-dependent adenoviral vector.

Construction of helper viruses and helper-dependent viruses can be achieved based on the guidance provided herein using techniques well known in the art, such as those described by Ausubel, *Current Protocols in Molecular Biology*, John Wiley, 1987-1998, and Sambrook, *et al.*, in *Molecular Cloning, A laboratory Manual*, 2nd Edition, Cold Spring Harbor Laboratory Press, 1989, both of which are hereby incorporated by reference herein. Additional descriptions of useful techniques are provided, for example, in the Example section provided below and references such as Mitani, *et al.*, *Proc. Natl. Acad. Sci. USA* 92:3854-3858, 1995; Fisher, *et al.*, *Virology* 217:11-22, 1996; Kochanek, *et al.*, *Proc. Natl. Acad. Sci. USA* 93: 5731-5736, 1996; Parks, *et al.*, *Proc. Natl. Acad. Sci. USA* 93:13565-13570, 1996; Parks, *et al.*, *J. Virology* 71(4):3293-3298, 1997; and Schiedner, *et al.*, *Nature Genetics* 18:180-183, 1998; each of which are hereby incorporated by reference herein.

The helper-dependent adenoviral vector and the helper virus can be generated in a wide range of host cells able to be infected by adenovirus. Examples of host cells include 293 cells, 911 cells and PERC.6 cells (WO 97/00326 hereby incorporated by reference herein). The employed cell can be constructed to provide for useful proteins such as viral proteins or a recombinase.

HELPER VIRUS

Helper virus elements useful for facilitating the production and isolation of helper-dependent viruses include a low homology excisable packaging signal cassette and a substituted E3 region. Preferred helper viruses contain both a low homology excisable packaging signal and a substituted E3 gene.

5

Low Homology Excisable Packaging Signal Cassette

The low homology excisable packaging signal cassette contains a modified packaging sequence which fulfills the role of an adenovirus packaging signal, has a low homology to the Adenovirus packaging signal, and is flanked by a 3' and 5' recombinase recognition sequence. Preferably, the modified packaging signal is less efficient than a wild type packaging signal. Advantages of such low homology excisable packaging signals include: (1) less recombination between the helper virus and helper-dependent adenoviral vector; and (2) when a less efficient packaging signal is used, less packaging of those helper viruses which escaped excision of the packaging signal.

15

Recombinase recognition sequences recombine when acted on by a recombinase that recognizes the sequences, resulting in the excision and circularisation of intervening nucleic acid. The recognition sequences should be sufficiently far apart, at least about 180 bp, to allow for the proper positioning of nucleic acid segments being recombined. Preferably, the recombinase recognition sequence is *loxP* or *frt*.

20

The *loxP* recognition sites can be positioned 3' and 5' of the packaging signal to allow for excision by Cre recombinase. Cells, such as 293 *cre* cells, stably expressing Cre recombinase can efficiently remove nucleic acid located between *loxP* recognition sites. (See, Parks, *et al.*, *Proc. Natl. Acad. Sci.* 93:13565-13570, 1996 and Parks, *et al.*, *J. Virology* 71(4):3293-3298, 1997, both hereby incorporated by reference herein).

25

The *frt* recognition sites can be positioned 3' and 5' of the packaging signal to allow for excision by Flp recombinase. Cells stably expressing Flp recombinase can efficiently remove nucleic acid located between *frt* recognition sites. Flp-mediated gene modifications are described in U.S. Patent No. 5,564,182, hereby incorporated by reference herein.

30

Homology between a helper and the helper-dependent adenoviral vector encourages recombination events between the two, resulting in unwanted

changes in the structure of the helper-dependent adenoviral vector or the helper virus, and leading to an increased contamination by helper virus. For example, if the first of a *loxP* site flanking the packaging signal is removed by homologous recombination within the packaging signal, the resulting helper virus escapes selection in 293 cre cells because the Cre recombinase is no longer able to excise the packaging signal.

Sequences for different low homology excisable packaging signal cassettes can readily be designed taking into account recombinase recognition sequences and adenovirus wild-type packaging signal sequences using the guidance provided herein. The provided guidance focuses on the adenovirus serotype 5 packaging signal to illustrate the production of a modified packaging signal. Such guidance is applicable to producing other modified packaging signals taking into account other adenovirus serotypes.

The wild-type packaging signal of adenovirus serotype 5 is formed by at least seven functional units called A repeats, which are located between nt 230 and nt 380 of the genome. The A elements have the consensus sequence ATTTGN δ GC (Schmid *et al*, 1997 *J. Virol.* 71:3375-3384, which is hereby incorporated by reference). The modified packaging signal of this invention preferably comprises less packaging elements than the wild-type (which has seven elements), preferably from about two to six elements, and more preferably from three to five elements. In a preferred embodiment, the modified packaging signal contains only four out of the seven original packaging elements (elements AI to AIV). The four elements are preferably in a modified form with two strong elements (AI and AII) present. The position of the elements relative to each other may be changed, but in preferred embodiments, it is maintained.

In order to reduce contiguous sequence homology, the eight ambiguous nucleotides of the consensus sequence (ATTTGN δ GC; SEQ. ID. NO. 1) within each A element are preferably replaced by sequences taken from a different A element. For example, the eight nucleotides within AI were replaced by those from AV; and the eight nucleotides within AII were replaced by those from AVI. In addition, a new element was created between AII and AIII starting 21 bp after AII, by changing the existing nucleotides to the consensus sequence. Two more nucleotides were exchanged within AIV: ATTTTGTGTT (SEQ. ID. NO. 2) was changed to ATTTTGTTGT (SEQ. ID. NO. 3). One embodiment of a synthetic packaging signal is given in SEQ. ID. NO. 4.

Desired nucleic acid sequences can be produced using different techniques including those involving the creation of nucleic acid mutations and nucleic acid synthesis techniques. Examples of such techniques are provided in Ausubel, *Current Protocols in Molecular Biology*, John Wiley, 1987-1998, and
5 Sambrook, *et al.*, in *Molecular Cloning, A laboratory Manual*, 2nd Edition, Cold Spring Harbor Laboratory Press, 1989, both of which are hereby incorporated by reference herein.

E3 Insertion

10 Homologous recombination between viral sequences (E1) present in the cell lines used for propagation of either helper virus (lacking E1) or helper-dependent adenoviral vector, and the helper virus itself, has resulted in the generation of wild-type viruses. Homologous recombination results in the insertion of the E1
15 region into an E1-deleted virus genome. To avoid this problem, in preferred embodiments of this inventions, the helper virus also contains an insertion element. Preferably, the insertion element does not contain a promoter or encode for a protein when present in the helper virus. More preferably, the insertion element is non-coding repeat-free human DNA devoid of splice signals.

Parks, *et al.*, *Proc. Natl. Acad. Sci. USA* 93:13565-13570, 1996,
20 describe the use of an E3 region containing an insert encoding for the ampicillin gene, the bacterial origin of replication, a CMV promoter and the luciferase gene (AdLC8BHG10luc). The insert protected against the production of viable viruses having a wild type E1 produced by homologous recombination. However, the insert impairs virus generation, most likely by interference with fiber gene expression. The
25 resulting level of viral generation is disfavored for the production of viral stock.

The use of non-coding sequences as an E3 insertion element, preferably human non-gene sequences, provides for more efficient viral generation than that observed with E3 insertional elements encoding for bacterial protein described by Parks, *et al.*, *Proc. Natl. Acad. Sci. USA* 93:13565-13570, 1996 (see
30 Table 1, *infra*.) The use of intron sequences are preferred because such sequences do not contain hidden splice signals that could interfere with correct splicing of the fiber message.

The size of the insertion element (preferably between about 2800 and 3500 bp, more preferably about 2900 bp) is chosen to avoid the generation of

potentially harmful wild-type viruses while maintaining a virus size close to wild-type size. Starting with a helper virus having a deleted E1, if homologous recombination occurs resulting in insertion of the E1 region, the recombined adenoviral vector which also contains the E3 insertion element would have a genome size of approximately
5 38350 bp (108.1% of wild-type size). Because this size exceeds the packaging limit of Adenovirus 5, no viable virus would be produced.

The insertion element may be from virtually any source. In a preferred embodiment, a 2900 bp insert was taken from human chromosome 11q13, starting 13340 bp upstream of STS marker 11S1337 (available from Genethon) and inserted
10 into the Xba I site at position (nt) 28593 (referring to Ad5 wild-type).

The insertion element and its insertion point in E3 should be chosen to avoid any reduction in synthesis of the fiber protein. Such a reduction is likely to occur if the E3 region is modified because synthesis of fiber protein is dependent on correct splicing of RNA from signals located in this region. We demonstrate that
15 viruses containing the insert grow to titers similar to those obtained with viruses containing an unchanged E3 region.

HELPER-DEPENDENT ADENOVIRAL VECTOR

Helper-dependent adenoviral vector elements useful for facilitating the
20 production and stability of helper-dependent adenoviral vectors include an E4 non-coding segment that confers a selective advantage to a virus and stuffer region(s) that provide a helper-dependent adenoviral vector with a GC content of about 50% to about 60%. Preferably, the helper-dependent adenoviral vector contains an E4 non-coding segment conferring a selective advantage to the vector; and stuffer region(s)
25 providing the vector with a GC content of about 50% to about 60% and an overall size of about 28 kb to about 36 kb. In a preferred embodiment, there are no bacterial plasmid-based sequences (such as an origin of replication or bacterial marker genes) present in the helper-dependent adenoviral vector.

30 Stuffer DNA

Stuffer sequences are used to provide a helper-dependent adenoviral vector with an overall size of about 28 kb to about 36 kb. A helper-dependent adenoviral vector containing only a heterologous expression cassette, adenoviral ITR's and a packaging signal may be quite small, in general about 5-10 kb, depending

on the size of the heterologous expression cassette. Because such a small virus does not package efficiently stuffer sequences are added to provide a final size of about 28 kb to about 36 kb. Preferably, the stuffer provides the vector with a GC content of about 50% to about 60%.

5 Stuffer sequences used in gene therapy applications should not contain active genes or be recognized as foreign in a human cell. A preferred stuffer sequence replicates as well as the adenovirus itself, is not recognized by the host as foreign DNA and does not lead to chromosomal integration of the transgene.

10 Genomic DNA of mammalian origin, or even more preferably of human origin, are preferred stuffer sequences for gene therapy applications to avoid eliciting an immune response. However, stuffer based on human genomic DNA is identical to chromosomal sequences in a human target cell and could allow for insertion of the heterologous expression cassette into the host genome by homologous recombination. This phenomenon could have a potentially dangerous effect and
15 should be prevented.

To prevent integration of the heterologous expression cassette into the host genome, the contiguous stuffer sequence can be interrupted, the individual fragments reversed and the expression units inserted at the specific breakpoints. (See for example, Figure 1.) Using this strategy, the transgene is not flanked by regions
20 which would support homologous recombination. Additionally, potentially unstable genetic elements such as retrovirus LTR as well as genomic repeats (*e.g.*, MIR, and ALU) should avoided or removed to prevent rearrangements during amplification.

Another consideration is the GC content of the virus. The GC content that stuffer sequences provide to a helper-dependent adenoviral vector has a major
25 impact on yield and purity of the virus. Competition experiments between different viruses as well as individual analysis revealed a correlation between the GC content of a virus and growth properties. The GC content is preferably between about 50% and about 60% and even more preferably between 52% and 57% which is close to the GC of wild-type adenovirus. Without being limited to any particular theory,
30 advantageous growth properties are believed to be based on more efficient generation of a virus with a higher GC content.

Preferably, stuffer sequences are provided such that the heterologous gene cassette(s) that are present are at least about 1.0 kb from the helper-dependent viral vector 3' end and 5' end, more preferably at least about 2.0 kb, and more

preferably at least about 4.0 kb. Additionally, high AT rich segments such as those provided in the hpvt fragment (59% AT) may be unfavorable even with a viral vector having an overall high GC content. A "high" AT is above about 55% AT.

5 E4 Non-coding Segment

 The E4 region of adenovirus type 5 occupies about 3000 bp at the right end of the genome. Transcription from the E4 promoter is directed from the right end toward the center. Seven polypeptides are produced from open reading frames (ORFs) located within the region with the first ORF starting at nucleotide (nt) 406
10 from the right end of the virus.

 It was observed that the E4 region is sometimes transferred from a helper to a helper-dependent adenoviral vector as a result of homologous recombination. This recombined helper-dependent adenoviral vector generally is favored in co-cultivation with the non-recombined helper-dependent adenoviral
15 vector and dominates in vector preparations. Despite its strength, this recombined helper-dependent adenoviral vector cannot be used as a helper-dependent vector for gene therapy because it contains functional E4 genes. Expression of those E4 genes in a host cause an immune rejection of vector infected cells by the host.

 In accordance with this invention, it has been determined that a
20 noncoding segment of the E4 region confers a selective advantage when a virus is co-cultivated with a virus lacking the same element. The entire E4 region is not needed.

 From adenovirus serotype 5, the part of E4 between the 3' ITR (starting 102 nt from the right end) and nt 405 which does not contain coding sequence is the segment that should be preferably included into the helper-dependent vector and is a
25 non-coding E4 segment. E4 sequences located left of this element have no additional impact on virus strength. This is of importance because all genes (open reading frames, orfs) are located in the left part of the E4 region. Thus, another aspect of this invention is a *cis*-acting sequence located next to the 3' ITR (right virus end) which is preferably incorporated into the helper-dependent adenoviral vector because it
30 supports amplification of the helper-dependent adenoviral vector.

 Competition experiments were designed to determine the location of the beneficial element. Helper-dependent adenoviral vectors based on the vector stk120 (which contains only a 5' ITR, a 3' ITR and packaging signal as its viral-based nucleotides, as described in Morsy *et al.*, 1998, *Proc. Natl. Acad. Sci. USA* 95:7866-

7871, which is hereby incorporated by reference) were constructed containing either the 3' ITR only (stk120gfp), the 3' ITR and noncoding sequence of the E4 region until nt 400 from the right end (stk120gfp-E4promoter) or the 3' ITR and the complete E4 region to nt 3115 from the right end (stk120gfp-E4). Pairs of these vectors were
5 cotransfected at a 1:1 ratio to originate rescues of two competing helper-dependent adenoviral vectors.

After passage seven, viruses from at least three independent rescue procedures were purified, and DNA was extracted and analyzed by radioactively labeled restriction digests. The two competing helper-dependent adenoviral vectors
10 have an almost identical restriction pattern. The only difference is in the size of one fragment, called the discriminating fragment. The discriminating fragment was found 5-10 fold stronger for stk120gfp-E4-promoter than for stk120gfp, and at least 4 fold stronger for stk120gfp-E4 than for stk120gfp (see Figures 2a and 2b). However, the discriminating fragments are of equal strength for stk120gfp-E4-promoter and
15 stk120gfp-E4 in each rescue (see Figure 2c). Therefore an E4 element supporting helper-dependent adenoviral vector growth is present in stk120gfp-E4 promoter and stk120gfp-E4 and must be entirely located in the right part of the E4 region (between nucleotide -400 from the right end and the 3' ITR).

While not wishing to be bound by theory, it appears that there are two
20 possible explanations for the selective advantage to the E4 segment-containing adenoviral vector: the E4 element allows for more effective replication of the virus or it confers improved packaging and stability. Our results indicate that the E4 segment does not effect replication. If infection is started with equal amounts of helper and helper-dependent adenoviral vectors, the ratio between helper and helper-dependent
25 genomes inside the cell at the end of the production cycle (48 hours) is identical. However, 3-fold more helper-dependent adenoviral vectors were packaged for the genome containing the E4 segment compared to contaminating helper genomes. Thus, it appears that the E4 segment supports effective packaging.

30 EXAMPLES

Examples are provided below to further illustrate different features and advantages of the present invention. The examples also illustrate useful methodology for practicing the invention. These examples do not limit the claimed invention.

EXAMPLE 1

General Materials And MethodsConstruction Of Plasmids

5 Cloning techniques described in J Sambrook, E.F. Fritsch and T. Maniatis, 1989 *Molecular Cloning*, Cold Spring Harbor Laboratory Press, which is hereby incorporated by reference, were used.

To generate plasmids containing the entire Ad5 genome, homologous recombination in *E. coli* based on the recF pathway was used. Competent cells of *E.*
10 *coli* strain BJ 5183 were transformed with a plasmid vector containing the adenovirus genome and a fragment from the shuttle vector. This fragment contains the sequence to be inserted flanked on both sides by sequences homologous to the vector. In order to select against re-transformation of the vector without modification, it was linearized close to the insertion point. DNA isolated from individual colonies of
15 BJ5183 was retransformed into the strain XL2 (Stratagene) and high yield plasmid preparations were obtained. Correct clones were determined by restriction analysis and sequencing.

Virus Rescue

20 10 μ g helper plasmid was digested with PacI and transfected by calcium phosphate coprecipitation into a 6 cm dish of subconfluent 293 cells. Directly after transfection cells were overlaid with α MEM/ 10% FCS/ 0.8 % Seaplaque agarose. 20-100 plaques appeared after 7-9 days. Plaques were taken 12 days after transfection and used to reinfect cells. After two intermediate passages the
25 cell lysate was used to infect NUNC-cell factories (NUNC). Cells were harvested 48 hours after infection and virus was released by three freeze/thaw steps (-70°C /37°C). After treatment of the cleared lysate with BENZONASE nuclease, viruses were purified by ultracentrifugation on a CsCl gradient. After one step gradient ($q=1.5 / 1.35 / 1.25 \text{ g/cm}^3$) the viruses were loaded on a continuous gradient formed from
30 a CsCl solution of $q=1.35 \text{ g/cm}^3$. Viruses were dialyzed against 10mM Tris pH 8.0, 10mM MgCl₂ and 10% glycerol and stored in aliquots at -70°C. Helper-dependent adenoviral vectors were purified using substantially the same procedure. For analytical purposes virus obtained from two 15 cm dishes was purified. Virus DNA was cut with Hind III or other restriction endonucleases, restriction fragments were

labeled using $\alpha P^{33}dATP$ or $dCTP$ and Klenow polymerase and separated in 0.7% agarose gels. The gel was dried and exposed to film.

Determination Of Particle Titer

- 5 10 μ l of virus suspension was added to 90 μ l PBS/ 0.1% SDS and incubated at 50°C for 20 minutes. OD₂₆₀ is determined spectrophotometrically and the concentration of virus particles was calculated based on the equation 1 OD₂₆₀ = 1.1x10¹² particles.

10 Extraction Of Viral DNA and Restriction Digest

- 100 μ l of virus suspension were added to 100 μ l lysis solution (10mM Tris-HCl pH 7.5, 10mM EDTA, 0.5% SDS, 0.05% pronase) under constant vortex. After an incubation at 37°C for at least 2 hours followed by addition of 100 μ l TE the lysate was extracted 1x with Tris-HCl saturated phenol. Virus DNA was precipitated
15 with two volumes of ethanol, washed extensively with 70% ethanol and resuspended in TE. Virus DNA was cut with Hind III or other restriction endonucleases, restriction fragments were labeled using $\alpha P^{33}dATP$ or $dCTP$ and Klenow polymerase and separated in 0.7% agarose gels. The gel was dried and exposed to film.

20 Extraction of Cellular DNA From Infected Cells

- 10⁶ cells are lysed in 10mM Tris-HCl pH 7.5, 10mM EDTA, 0.5% SDS and treated with 100 μ g/ml Proteinase K overnight at 58°C. The lysate was extracted with Tris saturated phenol, phenol/chloroform/ isoamylalcohol (24:24:1) and chloroform, precipitated with two volumes of ethanol, washed extensively with 70%
25 ethanol and resuspended in TE.

End Point Dilution Assay

- 96 well plates were seeded with 1x10⁴ 293 cells in 100 μ l per well 24 hours before the assay. Each virus stock was diluted 10³ to 10⁸- fold. Dilutions were
30 used to infect 24 wells each (50 μ l/ well). Positive wells were scored by cytopathic effect after 14 days and the virus titer was calculated based on two dilutions with intermediate numbers of positive wells taking the dilution factor into account.

Quantitative PCR

Real time quantitative PCR (ABI PRISM 7700) was used to determine the relative amounts of helper and helper-dependent adenoviral vector. Two specific target sequences were selected present either in all helper viruses (Ad5 sequences from 11358-11456) or in all helper-dependent adenoviral vectors based on stk120 (bp 12037-12176) (Morsy *et al.*, 1998, *Proc. Natl. Acad. Sci. USA* 95:7866-7871). A plasmid containing both target sequences was constructed and used as one standard for both amplifications. Within each target sequence, a set of forward primer, reverse primer and probe (which is located between the primers and contains a fluorogenic reporter and a quencher) have been selected. Separation of quencher and reporter generated a fluorogenic signal during the logarithmic phase of PCR amplification which was plotted versus cycle number. The plot was used to calculate template concentration based on the software of the 7700 system (ABI).

EXAMPLE 2

Construction Of Helper Viruses

Helper viruses for the helper-dependent system were made based on Ad5. In order to obtain specific helper functions, the plasmid pAdE1-E3+ (containing the entire Ad5 genome with a complete deletion of the E1 region) was modified in two regions: i) between the left ITR and the promoter of protein IX; and ii) in the E3 region.

To introduce changes at the left end of the genome a shuttle vector "plox Δ pack" was constructed. It is based on pUC (Boehringer Mannheim) and contains the following elements:

- A) nt 1-195 of Ad 5 linked to a unique Pac I site immediately before nt 1;
- B) a *loxP* site starting directly after nt 195 of Ad5;
- C) a synthetic packaging signal contained in the following sequence:
5'- GTACACAGGA AGTGACTTTT AACGCGCGGT TTGTTACGGA
TGTTGTAGTA AATTTGTCTA GGGCCGAGTA AGATTTGACC
GTTTACGCGG GGA CTTTGAA TAAGAGCGAG TGAAATCTGA
ATAATTTTGT TGTACTCATA GCGCGTAATC TCTAGACG-3' (SEQ.ID.NO:4).
- D) a second *loxP* site;
- E) a linker containing cloning sites; and

F) 2247 nucleotides of Ad5 sequence starting from nt 3534 (according to wild-type adenovirus numbering).

For comparison, a second shuttle vector, "plopack", containing the wild type packaging signal instead of the synthetic one was constructed. It contained:

5 A) nt 1-195 of Ad 5 linked to a unique Pac I site immediately before nt 1;

B) a *loxP* site starting directly after nt 195 of Ad5;

C) Elements AI-AIV of the wild-type packaging signal contained in the following sequence:

10 5'-GTACA CAGGAAGTGA CAATTTTCGC GCGGTTTTAG
GCGGATGTTGTAGT AAATTT GGGCGTAACC GAGTAAGATT
TGGCCATTTT CGCGGGAAACTGAA TAAGA GGAAGTGAAA
TCTGAATAAT TTTGTGTTAC TCATAGCGCGTAATCTC TAGACG-3' (SEQ.
ID. NO. 5)

15 D) a second lox P site;

E) a linker containing cloning sites; and

F) 2247 nucleotides of Ad5 sequence starting from nt 3534 (according to wild-type adenovirus numbering).

20 The second shuttle vector was generated by cloning the 4.9kb *XhoI* fragment from Ad5 wild-type (Ad5wt) DNA (nt 24797-29791) into pUC. A 2900 bp fragment of human DNA located on chromosome 11q13 starting 13340 bp upstream of STS marker 11S1337 (Genethon) was obtained by PCR and inserted into the *XbaI* site at nt 28593 (Ad5wt) in either orientation. This fragment contains part of inton 2 of the human LRP5 gene.

25 A strategy based on homologous recombination between plasmids in *E. coli* (Chartier, *et al.*, 1996 *J. Virol* 70: 4805-4810, which is hereby incorporated by reference) was used to introduce those sequence elements into plasmid pAd5E1-E3+ containing a complete Ad5 genome with a deletion of E1. The procedure allowed for insertion of those elements at exactly the positions described for the shuttle vectors.

30 The flanking Ad5 derived sequences were completely maintained. The following plasmids for new helper viruses were constructed:

pAdlspLI1 = Helper 14 -- has synthetic packaging signal, intron of LRP5 in E3;

pAdlpLI1 = Helper 1 -- has wild type packaging signal, intron of LRP5 in E3;

pAdlpE3 = Helper 11- -has wild type packaging signal, wild type E3.

Plasmids containing the Ad 5 genome with the described modifications were digested with *PacI* to release the virus genome from the plasmid and transfected into 293 cells by calcium phosphate coprecipitation. Resulting virus plaques were used to reinfect cells. Amplified virus stocks were characterized by restriction and sequence analysis.

1. Growth of helper viruses in 293 and 293cre cells

Different helper viruses as well as a E1 deficient adenovirus (dl70-3) (Bett et al. *Proc. Natl. Acad. Sci. USA* 91:8902-8906 which is hereby incorporated by reference) taken as control, were grown in 10 layer NUNC cell factories, purified in CsCl Gradients and dialyzed. The virus concentration (particle concentration) was determined by spectro-photometric analysis (OD₂₆₀). 1.5 x 10⁶ cells (293 or 293cre) were infected in a 6 cm plate with 100 particles per cell. CPE was apparent in all plates after 48 hours. Cells and medium were harvested, frozen and thawed 3x and the titer was determined in an end point dilution assay. The number of infectious viruses produced per cell was calculated and the results are shown in Table 1.

Table 1

	dl70-3 Ad without lox	AdLC8BHG10luc	AdlpLI1 Helper 1 Wt. Ψ	AdlpLI1 Helper 14 Synth. Ψ
293cells	1200*	27	840	610
293cre4 cells	1400	1.2	1.3	0.8

*infectious units per cell

The amplification rate for all new helper viruses exceeds that of AdLC8BHG10luc (Parks et al., 1996, *Proc. Natl. Acad. Sci. USA* 93:13565-13570, which is hereby incorporated by reference) by at least 20 fold. The growth of the helper viruses in contrast to dl70-3 is well controlled in 293cre4 cells. These cells release only 1-2 infectious particles/cell. The rate of suppression is therefore

dramatically increased for the new helper viruses. Viruses without or with the insert in E3 in either orientation do not differ much in virus yields; thus, the chosen E3 insert did not have a major impact on virus growth.

5 2. Amplification of helper-dependent adenoviral vectors over a single passage

 A characterized stock (known particle titer and helper contamination)
of a helper-dependent adenoviral vector containing the human OTC gene was used to
infect 293*cre* cells. Cells were infected with 700 particles of the helper-dependent
adenoviral vector and 100 particles of one of the helper viruses per cell. After 48
10 hours cells were collected, virus was released through 3 freeze/thaw steps and purified
by Cs banding. Particle titers were determined by spectrophotometric analysis
(OD₂₆₀). Virus DNA was extracted and analyzed by restriction digests (Figure 3).

 Overall virus yield was slightly increased for the new helper virus over helper
AdLC8BHG10luc. However, the ratio between DNA-containing particles and empty
15 particles is increased for the new helper. Since helper DNA is undetectable in the gel,
helper virus does not contribute significantly to the overall output.

 To determine helper contamination more accurately, quantitative PCR
based on TaqmanTM technology was used. Contamination was similar between the
helper AdLC8BHG10luc and the helper LPLI1 containing the original packaging
20 signal, whereas the amplification supported by helper LSPLI1 containing the modified
packaging signal resulted in a lower contamination (0.29%) (Table 2). This is
expected because yields of virus grown in 293 cells are lower than those of an
identical virus containing the normal packaging signal. Therefore helper virus which
would escape *cre*-selection should still have a selective disadvantage.

25

Table 2

	HD Virus Productivity	Helper Contamination		
		%cont.	St. Dev.	%RSD
AdLC8BHG10luc	4800	0.52%	0.1%	22%
Helper 1 (Wt. Ψ)	5300	0.42%	0.12%	28%
Helper 14 (Synth. Ψ)	8100	0.29%	0.06%	21%

3. Construction of modified backbones

5 Constructions are based on plasmid stk120 (Morsy *et al.*, 1998, *supra*) containing a CMV promoter driven gfp (green florescent protein) gene within the stuffer sequences at a unique *Swa*I site (stk120gfp). In order to incorporate the additional elements into the helper-dependent adenoviral vector, the 3' ITR (130 bp) was extended to 400 bp from the right end of the Ad5 genome. This fragment
10 contains the complete E4 promoter region. However, no adenovirus coding sequences are present. The resulting adenoviral vector is stk120gfp-E4-promoter.

 In a second construct, a 1353 bp fragment from the right end of stk120gfp containing part of the stuffer sequence and the 3' ITR was replaced by a 3115 bp fragment of Ad5 from a *Mun*I site to the end of the genome (stk120gfp-E4).
15 This fragment contains the complete functional E4 region.

4. Competition between helper-dependent adenoviral vectors

 Genomes for stk120gfp and stk120gfp-E4-promoter were released from the respective plasmids and cotransfected with a circular plasmid containing the
20 entire helper genome into 293 cre cells at a 1:1:2 ratio followed by infection with helper virus at moi 1.

 The mixture of viruses was amplified up to passage 7 where the viruses were Cs-purified. Five independent experiments were carried out. Virus DNA was extracted and analyzed for presence of the two virus types. In all
25 experiments a 5-10 fold excess of stk120gfp-E4-promoter was found (Figure 2a). This indicates the presence of a *cis*-acting sequence located next to the 3' ITR (right virus end) which supports helper-dependent adenoviral vector growth.

An identical experiment was carried out to determine whether the complete element is located within 400 bp from the right virus end or sequences further upstream contribute to this feature. *stk120gfp-E4-promoter* and *stk120gfp-E4* were cotransfected with the circular plasmid containing the entire helper genome (Helper 14) into 293 *cre* cells at a 1:1:2 ratio followed by infection with helper virus 36 hours after transfection at moi 1. After passage 7, viruses from three independent rescue procedures were purified, DNA was extracted, digested with *BsrGI* and labeled. The discriminating fragments have equal strength for each rescue (Figure 2c). Therefore, the cis-acting sequence supporting helper-dependent adenoviral vector growth is located between nucleotide -400 from the right end and the 3' ITR.

Example 3

Construction Of Helper-Dependent Adenoviral Backbone Containing Alternative Stuffer Sequences

The construction of the new helper-dependent adenoviral vectors (group A) began with the ligation of a 3498 bp fragment from pSTK120 (which comprised of the bacterial sequences origin of replication and ampicillin resistance gene, the right and left ITRs and the packaging signal) ligated with a multiple cloning site linker (SEQ. ID. NOs. 6 and 7) containing *HindIII* sticky ends and cloning sites *EagI*, *AscI*, *BglII*, *NotI*, *XbaI*, *MluI*, *EcoRI* and *KasI*. This clone is named "pITRF".

DNA from the region encompassed within HUMDXS455A (Genebank Accession # L31948) used as stuffer DNA was amplified from human genomic DNA in four segments of about 4.5kb each using the following pairs of primers:

- SEQ. ID NOs. 8 and 9 (PCR 1)
- SEQ. ID NOs. 10 and 11 (PCR2)
- SEQ. ID NOs. 12 and 13 (PCR3)
- SEQ. ID NOs. 14 and 15 (PCR4)

PCR1 product was Klenow filled in and then digested with *AscI*. It was ligated directly into the *AscI/EcoRV* of pITRF to produce "Construct 1".

PCR2 product was digested with *NotI/AscI* and cloned into the *NotI/AscI* site of Construct 1 to produce "Construct 2". PCR3 product and Construct 2 were digested with *MluI/NotI* (double digest) and ligated to produce "Construct 3". The PCR4 product was first cloned into a TA cloning vector.

To avoid inefficient direct cloning for very large plasmids homologous recombination in *E. coli* BJ 5183 was used to insert PCR4 and additional genomic fragments. The technique requires a small shuttle vector containing the region which overlaps with the insertion point in the large plasmid. This shuttle vector was
5 constructed as follows: 2 kb of the junction over PCR3 and vector sequence was amplified using Construct 3 DNA. The PCR fragment was Klenow filled in and then digested with SapI enzyme and ligated into PABS helper-dependent-3.0, a small plasmid vector (KanR) cut with EcoRI followed by a fill in reaction using Klenow and a digestion with SapI. Homologous recombination between Construct 3 and the
10 shuttle resulted in Construct 4 (C4).

The helper-dependent adenoviral vector part of C4 is 20kb in length and has 3 unique restriction enzyme sites to incorporate 3 different transgenes (AscI, NotI and SwaI).

As a next step, a CMV promoter driven gfp gene was inserted into the
15 NotI site of C4 by homologous recombination. For this purpose, a shuttle vector was built containing 2 kb of C4 overlapping the NotI site (pNot). The gfp expression unit was inserted by blunt end cloning into NotI and a fragment of this vector was used for homologous recombination with C4. Because a single NotI site is also present in the gfp expression unit, the gene can later be replaced by other genes using this restriction
20 site. The gfp gene is located at a junction between 2 inverted genomic fragments which prevents insertion into the host genome by homologous recombination which would be likely, if the gene was inserted within a contiguous sequence.

To incorporate additional stuffer sequences and extend the 3' ITR to include the E4 promoter, a second shuttle based on pABSHD-3 was constructed
25 containing 2 kb of C4 including the 3' ITR and unique KasI, SalI, BglII and HindIII sites for insertion. The E4 promoter was directly cloned as a Hind III/ SalI fragment into this shuttle vector.

Stuffer DNA fragments AFO, HSU, ER1, and ER2 (Figure 4), were amplified from human genomic DNA by PCR Expand Kit (Boehringer) and cloned
30 into the Topo-blunt vector (Invitrogene). AFO and HSU were inserted individually as SalI fragments, the ER1 and ER2 fragments were inserted sequentially after KasI/ BglII and BglII/ Sal digests. Recombination in *E. coli* BJ 5183 lead to the final backbones C4HSU (SEQ. ID. NO. 16), C4AFO (SEQ. ID. NO. 17), and C4ER (Figure 4).

Construction of the new helper-dependent backbones (group B) was started by replacement of a 8181bp fragment from stk120 by the gfp expression unit. For this purpose a shuttle vector was built based on pUC 18 containing 1kb 5' to the unique SmaI site and 1kb 3' to the unique EagI site in stk 120. The gfp expression unit was inserted into the Bam HI site of this shuttle as a BglII fragment. The replacement was carried out by homologous recombination in *E. coli* BJ 5183 between STK 120 linearized with EagI and a fragment of the shuttle vector. This step removes a part of the hprt region from the vector which was most frequently lost during adenoviral vector propagation. This vector "STKSEgfp" is 19.4 kb in length and misses about 10kb of genomic sequence required for propagation as a helper-dependent adenoviral vector.

Fragments AFO and HSU were cloned after SalI digestion into the Sal I site of shuttle vector pSH1-3ITR which contains part of the HUMDXS455A - fragment in STK120 and the 3' end of Ad5 starting from nt -400 containing the E4 promoter and the 3' ITR. The fragments ER1 and ER2 were inserted step by step into the same vector as SalI/BamHI and BamHI/BglII fragments. The 3 resulting different shuttle vectors were recombined with STKSEgfp linearized with PacI to generate HXAFO, HXHSU and HXER (Figure 4).

20

Example 4

Generation Of Helper-Dependent Adenoviral Vectors

Cells from subconfluent plates were washed with PBS, trypsinized and seeded to 6 well plates (Costar) at 3×10^5 cells per well in 1.5ml medium. Genomes for helper-dependent adenoviral vector were released from the respective plasmids by Pme I digestion and 2 μ g of DNA were cotransfected with a 2 μ g circular plasmid containing the entire helper genome per well. A 24 hour transfection was followed by infection with helper virus at 1 moi. Forty-eight hours later cells were lysed by 3 freeze/thaw cycles and the lysate was used to infect 10^6 293cre cells together with Helper virus at MOI 5. Cells were lysed 48 hours later.

30

The same procedure was repeated with increasing the number of cells until passage 6 at which two 15 cm plates are infected. At this stage helper-dependent adenoviral vector is isolated by CsCl banding and analyzed by restriction digestion. Passage 6 is also used to infect a large scale preparation (1x10 layer Nunc cell factory).

35

Example 5

Competition Between Helper-Dependent Adenoviral Vectors

Virus rescue in a competition experiment was performed using the procedures described above and, instead of DNA from a single vector plasmid, a
5 equimolar mixture of 3 vectors was used. Three mixed rescues were carried out: one with HXAFO, HXHSU and HXER; one with C4AFO, C4HSU and C4ER; and the third with HXHSU and C4HSU. No rearrangements were observed in either the individual or the mixed rescues.

As shown in Figure 5 there are only moderate differences between
10 backbones within the first 2 groups with a slight dominance of the C4HSU construct. However, in a competition experiment between HXHSU and C4HSU the latter clearly dominated. Thus, the remaining hprt sequences in HXHSU have an adverse effect on adenoviral vector propagation.

15

Example 6

Analysis Of Large Scale Preparations Of Helper-Dependent Adenoviral VectorsSTK120-EF1amEPO, C4AFO-EF1amEPO and C4HSU-EF1amEPO

To investigate the behavior of the new helper-dependent adenoviral
vector *in vivo*, the gfp gene was replaced in STK120, C4AFO and C4HSU by the in
20 mice immunologically inert mEPO gene. Replacement was carried out using the shuttle pNot by homologous recombination. Viruses were rescued as described above. Large scale amplifications were carried out in Nunc cell factories and the virus was purified as described above.

Since the size of the helper-dependent adenoviral vector and the helper
25 differ only slightly, the purification procedure does not remove helper virus from the preparation. Such a physical separation was not intended to allow for a replacement of gradient based by chromatography based purification in large scale. DNA was extracted from purified virus and the content of helper virus relative to the amount of helper-dependent adenoviral vector was determined by Taqman™ quantitative PCR.

30 We observed a correlation between the helper content and the GC content of the helper-dependent adenoviral vector (see Figure 6). The experiment with Construct C4HSU has the lowest helper content (about 0.18%) compared to a helper content of about 1% for stk120.

Example 7

Analysis Of The Replication Potential OfDifferent Helper-Dependent Adenoviral Vectors

5 Cells were infected with 100 part/cell of helper and 300 part/cell of the helper-dependent adenoviral vector. Three hours after infection 1/3 of the cells was harvested and washed with PBS. Another portion was harvested at 42 hours and a third fraction at 48 hours. Cellular DNA was extracted as described above and the relative amount of helper and helper-dependent adenoviral vector DNA was
10 determined by Taqman™ PCR (Figure 7).

Figure 7 shows that if the rescue is started with an excess of helper-dependent adenoviral vector, the ratio of helper-dependent/helper is maintained and even slightly increased for C4HSU, whereas it is strongly decreasing for STK120 and slightly decreasing for C4AFO. This shows that the nature of the backbone has a
15 strong impact on the speed of replication. Whereas the Adenovirus genome coevolved with the adenovirus polymerase, the helper-dependent adenoviral vector sequence was artificially selected. Therefore a less efficient replication is expected. However, the most important determinant seems to be the GC content of the helper-dependent adenoviral vector.

20

Example 8

Sequence Information

The different sequences for the SEQ. ID. NOs. referred to in the application are as follows:

25

SEQ. ID. NO. 1
ATTTGN₈GC

SEQ. ID. NO. 2
30 ATTTTGTGTT

SEQ. ID. NO. 3
ATTTTGTGT

SEQ. ID. NO. 4

GTACACAGGA AGTGACTTTT AACGCGCGGT TTGTTACGGA
TGTTGTAGTA AATTTGTCTA GGGCCGAGTA AGATTTGACC
GTTTACGCGG GGACTTTGAA TAAGAGCGAG TGAAATCTGA
5 ATAATTTTGT TGTACTCATA GCGCGTAATC TCTAGACG

SEQ. ID. NO. 5

GTACA CAGGAAGTGA CAATTTTCGC GCGGTTTTAG GCGGATGTTGTAGT
AAATTT GGGCGTAACC GAGTAAGATT TGGCCATTTT
10 CGCGGGAAAACTGAA TAAGA GGAAGTGAAA TCTGAATAAT
TTTGTGTTAC TCATAGCGCGTAATCTC TAGACG

SEQ. ID. NO. 6

AGCTCGGCCGATTATTGGCGCGCCAGATCTGCGGCCGCTTCTAGAAACGC
15 GTGAATTCGGCGCCA

SEQ. ID. NO. 7

AGCTTGGCGCCGAATTCACGCGTTTCTAGAAGCGGCCGCAGATCTGGCGC
GCCAATAATCGGCCG
20

SEQ. ID. NO. 8

ATTGGCGCGCCTTCTTTCTGGGATGATTCAGCATCAACTC

SEQ. ID. NO. 9

25 GATCGTCGGCCGCTTGGGTCATAGACTTCTTTGAGAACCAG

SEQ. ID. NO. 10

ATCAGTTAGCGGCCGCACAAGCTAAGATCACAAAGCTGTTT

SEQ. ID. NO. 11

30 TATGGCGCGCCGCTGACACCCAGCCTGGGTGCCGGTG

SEQ. ID. NO. 12

TCGACGCGTAGCGCTGTGTGGCCTTGGCAGTTTCCATAG

SEQ. ID. NO. 13

TCAGTAATGCGGCCGCGGGATCATTCCTGGACTCAGATTGTTCTG

5 SEQ. ID. NO. 14

TATTAAGGCGCCGGGCATGGGAGTGATCTCACCAACTCTGG

SEQ. ID. NO. 15

TCGACGCGTATTTAAATGTGCTGGAGTGTTGAGATACTGTAGTGGT

10

SEQ. ID. NO. 16*Helper Dependent vector C4HSU*

The sequence is shown without transgenes.

AAACATCATCAATAATATACCTTATTTTGGATTGAAGCCAATATGATAATG
15 AGGGGGTGGAGTTTGTGACGTGGCGCGGGGCGTGGGAACGGGGCGGGTG
ACGTAGTAGTGTGGCGGAAGTGTGATGTTGCAAGTGTGGCGGAACACATG
TAAGCGACGGATGTGGCAAAGTGACGTTTTTGGTGTGCGCCGGTGTACA
CAGGAAGTGACAATTTTCGCGCGGTTTTAGGCGGATGTTGTAGTAAATTT
GGGCGTAACCGAGTAAGATTTGGCCATTTTCGCGGGAAACTGAATAAGA
20 GGAAGTGAAATCTGAATAATTTTGTGTTACTCATAGCGCGTAATATTTGTC
TAGGGCCGCGGGGACTTTGACCGTTTACGTGGAGACTCGCCCAGGTGTTT
TTCTCAGGTGTTTTCCGCGTTCCGGGTCAAAGTTGGCGTTTTGATTCGGCC
GCTTGGGTCATAGACTTCTTTGAGAACCAGTTATAAGCTATGGTTTCTCTC
CACAGAAAAGCACTTATGGTGTCTCCCCCTTCCAGCCCACCAACATTTT
25 ACATCTAATTTGGGGGGGTTTTCTGGACCACTTAATACCCATCCATGGATC
TCATGTGAAGACTCCCCTGGCTTGAGAAATCACTGTCTTGTTGAAAATGG
GAACAAAGCTAAGTCAGATAGCTGGTTCATCAGCAATGACTTTGACCAAG
CCTGATCCCACCCTACCCACCCACCCAGTGACCACCCCCACAATGG
AGCACACAACCTCTAAACTGGTTTGTAGGTATGTGTGTGTGAACACGCTGA
30 GGAATCTGCAAAACCAAATGGTGAGTGCAAAACCAAACAGTCACGAGTA
AATCTCACAACAACACGTCCTGAGCTGCAGCCCTTGTTGAACTATACCC
CACTAGGGCCCCAAGATTTTAGGACTTGTGTGTGGGTGGGACCTCCCCTTT

CTATCATGCTTTAGAAAGACAGGGATTTACACCAGAATTGAACATATTGAAC
ATATGACCCATTTTTTTTCAGCCAAAGGCAATTAAAATAACTTCATACTTGA
TATCCATGTCAGCAAAAGCTGCAAAACGCAAATGGGTGGCTGCTAAGAGC
CCTGGTACCCTGACGAGCACACCAAGTGCTTAGCAACAGTGGTGTCCAAA
5 GGACCAGCTGGAAGCCTGCCTTGATGAGAAGTTGCTCTTCTTTCTACATGA
AGGAACACCTCTACTCTCCTGCTTTTAATACCTGAGCTGTGAGTGATCATC
TATGTCCATTAGCAAACATCCCAGTGGAGAAGGAAACACTCATACCCGAA
ATCTAAGCTACATAGTTGGAATCACTTCAACTTATTGCAATAAACACTTAC
TAAGCACCTATTGTGGGCAAGTCTTTGCAATGGATAATAGTTCAGTAGAT
10 ATTTTGATGTAATATTTGAAATAACAATAAAAATTGCCACCACTGAATTTA
TTGAGCATTTGCTGTGCTTTAGGCACTAACCCAGGTTCTTTAAATATTTGG
TCTTATTCGATCTGTATAAATAGCCATCTATGAGAAAGGGACTATTATTGC
CCTTATTTTACAAATGAGGCCAATGAGGCCCCAGAGAGGTTAACTAAGTTG
CCCAAAATCATACAGCCCACTAGTGGCAGAGCAGGATGCAAACCCAGGC
15 TTGCCTCGTTCCCAAGCCCACATGTCGTTTGCATTGGTGTGGAGGTGTGCA
TGTGTTTAGTCATTAGCATGTTATATGATAAGCAAGTTTTGAAACATAGAA
ACTTAAAATGTGCCATTAAGAAAAGTACAGGCAAGGTTTTCCAAGGGGAG
GTGTGGACCTCCGGACAAATTTTAAAGAACTAATTATAAATACTTAAAAA
TGGAATAAGAAGACAACCTAACTACCTGAACAGTTTTAGAGATGACTCA
20 TGCCCACCCTCTAAAACCCAAACAAAAACAACAAAGTCAAGAAAACCCA
TGAAATCTTAGCAAGCGATTTCTATGTACTTGTGAAAAGGATTTCTTTACC
ATTCTAATGGGATTTATGCCAACCATAGAGGGCTCAGTGCCCCTCCCATG
GGGTGGTTAGTGAGTACAGAGCTGAGCTCACCGGCCATCTGCAGCTTCAT
GTTATCAAGCTCCAGTTTGTCTTGGAGCAAGGTTATCTGGGACATGAGC
25 AGAGGCATTGCTTTCTGCAATGGACAGTTCTTTCTGCCTGCATACCTAGCT
CCTTGATAACTTTAAATACCATTTTATAGCCACACTGGAGTTTTGAAGACC
TCAATATGCAAATATTACTCAGGTTCTGATTACTTGTCTGCTCCATGATAA
CACACTCTAAAAGCAATGAATGGGGCTTATTTGTAGAGAACTGAAGCATT
TTAAGCTTTTGCTCAGGAATCCCTGGTAGCTTCCTGTGACTTGCAAGATAT
30 TAGTGATGGGTCAAGAAACAGGACCCCCCATCAGCATAACATACGCAGTG

CCTCAGTAGTCCATCAGGCAGAAAAAACTGCAGATGGCACATGGGAAATG
ACCAGCGGCGGAAGATACCCCGACAGTGTGGGCAGTTCTATTTTCAGCAGC
AATCAAGAGGGGGCCTGGAGCCACTCAATCAAGTGGAGCAGGATGGGAG
CAAGCACTGTGCAGACCAATGCAATCCCCAGTTAACACAAAAAATAAATA
5 AAAGAGATGAGATTCAGTCTCTTGACTGTGACTGACTGGGAGCTTTATAG
CTGATGCTTGTGTCTTTTTCTCCATTTTATTTAATTAGGAAAAGAAATGCTT
ATCACACACTCTACGTGTGAGGTACAACCTCCACAGGAAAGGTTGTTAGG
AACATTTCAACTTCTAGAAGTTTCTAAACATAAGGTAAATCCATCTTTGTC
CTTGGGATCACTGCACATCTCAGAAAGGCAAATAAATCAGTAATTGGTGG
10 GCATAATTACTAGCTCATGGACTGACAAGGTCTACACTATTTTGAATCTCA
CAGAAGTAAGCCATGGGACAGATAGAGTCTGATAGTGGTGCCCCGTTTCC
TGGAGGTCACACTTACTCATCCCCCTGGACCCTGGGCTTCTCATGATTGTC
AGAGAGTTTGCTGGAACCAGGTCAGCCCAGTTTCCCTTCCCCTGAAAAAT
CCTCCAATGGCTCGCACCAAGACTAGAGATGCAAGTGCAAGCACATCCAC
15 CCTCTCAGCAGCCAGGTTTTGCGTTCCATAATGTCACGTACCCCCAGTCAT
ACCAATCTCCTTGGAGCTCTCCAGACAGGCTGCCATGTGTGGTCGGCCCTC
TGTGCTTGTGCTCCTTGGTTTGCCAAGCCTGGAATGCCCTTTCTCCATGAT
TGTA CTCTGGGAATTCTGTGTGTCTTTTCGAGATGAAGCTCCTCCACCCTGG
AAATTCTGCCTCTCTTTCCAGGTCCAGCTCCTGTGGCTGAATACCCTTGGC
20 TGA CTGAACATTCTTGCTGGGCCAAGTCTTAATCATCTCTAAATCCCTCTG
GTGCCCCGACAGTATCTAGCCCAGGCAAAGGGCTCAGCAAATACTTGCAA
AATTCAATAAACTTTACTATGTTACTAGCTTGGCATGATGTTCTAGGCACT
TGGGAGATTATAATCTGGTGAGCATTGTGCTAGTCTTTTTTTTGGGATAAA
CACCAATGATAAACTGAGCCTTTACTCTCATTTTAGCACATCTACTCATCT
25 CTTTTCAATGCTGGTGGGTTTATAAAACCCATCAGTAGAACAGAGAGTGA
TCATAATCATATGGTGAAAATAACATCAGCTAACATATTTACTGAACATG
CTTTAGTGTGCTGGGCACTGAACTCTGTGCACATGTGAATTTGAGATAGAT
TGTTCTAGCTAATAAGATGAGGAAGTGGAGATGGGCTTACTCAAGTCAC
ACAATAGCAAGATGGGGTACAGACAGAAACCAAGCTAGAAACCAAGCAA
30 CCCCTGGGTTTGGAAATGCATGGGCTCCTCCTCTGCACATGGCGAGGAGC

AGTCAGGTGCTCCTCCTTCTCTCATACTGAAATAACTCTGCACTTTTGGCT
ATTCTGTGACACTCTGTGCTATTCTGTAGCTCAAATGGCTCTGGTGCAAGG
AGCTCAGATTATGACTAACCCCATAGATATTCAGCTGCTTTGCAAGAAG
TGGATGAATGCTCTGGCTTTATAAATTATTGACTAGATTGGATATTGGCCC
5 AATCTCCACTCTGGTGATTTCGGAAGGAGGCATATGCACATTTGCAAGGTT
ATAGTAGTGCACTCTAATTCAGTGGCTTCTGAGAGCTTGTAGGTTTCTTG
ACTTAATCATTCTGGAATTAAGGTAATGGATCCTCAACACCTTTTCTTTCC
CTGGCCTTTACTAACCATGTAACAGAAATGAGCAGAGAAAACCCAAGAA
AGCGAACTGGGAGACTTGATGAGTGTGTCAAAGATGCTCAGAGTCCAAGGT
10 CCTCTGTGGCTCACTGACTTCAGAAGCCAACCTCCGTTGTTCAAGTCACTT
GTGAGGTTACTATGCTAGAGCACTAAATATTATCCAGATAGCCCAAAGAG
GTGAAGGCAGACATGTGGAAGAACCTGGATTTTTGGGCAAGTTGATGAAT
GCCTCCTGCCCCATATCAAAGAGAGGTGATAGGAGACAACCTTTGTGAATT
TGAAATAATGCACCGGGGAAACAGGAAAACGTAATGTAAGTCACGCTTCT
15 TGGCTTTTTTCTTGCCATTACCCTACTTGGCCAAGTGCAAATGGGATTTC
ATATATCATAAGTATGCATCTATTAATAACAATGCAAGAAAGCTGTACAA
CTGAAGTCTGAGATTTTGTAAAGAACAGAATCTCTGTAAGCATCACCATC
CAACAGAACTTCCTGAGTTGATGCTGAATCATCCCAGAAAGAAGGCGCGC
CAGCTCTGCAGGATCTTCAAGTCTGGGGTGCCACCAGCAAGCGACGGTCC
20 TCCATGGGCTCTTCACCTTACGGCAGTGTCCAGAGGCACCGCCAGTCCTCT
GCTCCTATGCTGGTCCTGCTGTCCCTGGCAAAGGAGCCAGAGCATTCTCT
CCAGGCCTCCCGAGGAGGCTGCTTCCTTTGTTTTGCAGATGGAGGCTCCCA
TCCTTTGTTCTGAATCAATGTGCTCCAAAGATAAGCCCCAAGAAAACAGT
TGTTGCCTTTTGACACTGACAATTAGAATCGTTGGAAAATGGAGAAAACA
25 GGAAATGGCAAATGGTTTCAGTGACCAGGAGGAAACCGTGCCTGAAAGTT
GCTGCTTAGTGACTGGGACACTCGCTTTCTGCTCTCTTATGAAGGACAGCC
TAGGCCGTGTGGCCTTTTATAACAAAGCTATGAAGGGGTCGTCAAATTT
TCTAGGGCTGCAACTGTGGCACTACGTCCTGTTGTGCCAGGTGACACTGA
CAAGCAGCACTGAGTTCTATGCAAGCCCAGGTGTGCTTCTCTCATGGTGA
30 CCCCCAGAGAACTAAGGCCAGCTCTTCCTCTGTACACCCCTCCAGCC

CCCACTGTCAGACAAGGGACCACATTCACAGACAGTCTCAGCCAAGATGG
CAACCTTGGAAGTCCTGGGGATGCCTTTCTAGAAGCTTTAGGCTGACGCC
AGGGAGCTGTAAAGCCCCCACCTGTCCCTGGGGGTTGTGTGCTGGGCAGG
AGAGAGGGGAAAGAGCCCCAACTCCACCACTGCCTGCTGGCACAACCTGA
5 GCCCACGCCAGGCACTGCACCAGCTTCACTCACCAGCTACCACATGCTAA
TGTCTCCTGTTTACCCAGGTGCAATCACCTGAGTGCTGCTTTCTCACCCT
GCATGGGAAGAACGATCTCATTACCAATTCAACCACTTAAGCAGGCCAAAA
GGACTCCAACGGGGGAAAGGCAGAGGACAAACGATTCGGGGACTGAGAC
AGATGCCCCACAAAATGTTCCCTGGATTTTAGCCGGATTCGGGCTGAAGTTC
10 CTTGCCCTATGAGCCCTTGGAGGAGAGCATCCTGACTCAGACAAGAATTC
AAATGACAAGACTTCTGGGGGATGGCAGTGAATGGATCCAAAGCCTTGTT
TTATAAAGAATGTGGAGGAAGGGAGGAAGGAAGGGAGGGAAGGGGAGA
CTGAGAGGGAGAGAGAATAAAGAAGAGGAAGAGAAGGAAGAGGAGGAG
GAGGAAGGGGAGGGGAGAGGAAAAAGGGGAGGGTAGAGGAGGAGGGAG
15 AAGGGGAAGGGAGGACAGAAGAGGGGAGGGAGCGGGGAGATTATCTAC
CCCTCCATGCTAGGGAACCTCCTGTCTCCTTTTCCAGCAGACACTGCTCCAT
GAGTGCCCCCACAGAGAGCCAGCCTGCCACACAGCAGAAGGTGCTCAGC
TATCTGTCCCAAAGGTTGAGCACTAGTGTATACCACAAGGAAAAAACCT
TCTTTTGTACTGTTTGGAAGCCAGAATTAGAAACAACCTCGTTCACCCTTCT
20 GACTTTCCTTTTGCAATGACAGGAAGATAGTATACTGCATGGGACCACCA
CACATTCAAGATCTGAGTTGTGGGGACCCCCACCCCCACATGGGTGGAGC
CCCCTGAGACCCGTGGGCTCTGACTGGCCATGGCCCATAGATAGTGACAA
CATACACACTGCTATCTGAAGCCTAGTTGGAGGACGGATGGATGCTTTTA
GGAAAAAATTCCCTTTAGTCGTCCCTCAGAACTACAACGGCTTCCTAAA
25 ATGAATGGCAGGCAAGTTTTGATTGTTTCCTAAGTTCCCCCTAACTCTCAG
AGCCTGTGACTCAGCTTCTCCAGAATGCCTGCAGGGAGGCAGGACCACTG
GGAGTAATGAGCACGAATTGGGCCACTTCTCCCAGTGGACTGTTAACGCC
GAGATCTGGGGTGGAACCCACTGAACACCTGCCCTTTTGC GTTGCCAGAG
ATGCAGCCAGAGGGATACGGGCCACAGTAACTCTGGGGCCCCTGAACCTA
30 CCTTCCCAAGACCCTGACGGGAAGGTATTGCTTACATGACAGTCCAGTCA

GTCCCTCTGTCACTCAGTTAACAACATGGTCATGGGTCTACTATGGGCCA
GGCTCTGTTCCGGGCATGGGGATAACACCGTGTGTCTCTACTCTTATGGAC
ACAAGTAAATCAATCAAAGCAGATAATGTTCTAATGACCACCCAGTAGGG
TGAAGGGATGGAGGGTGGGCAGGTGCTGGCAGAAACCTAAGGTCAAGTC
5 ATGGGACCCCAATGGGGAACTGGAACTGCCACACCATTTTCTATGTTAG
GGGTCTTTCCTTGCCCTTCTAGAGACCTCACCCATGGTAATGATGCTGCCC
TCAGGTTTGAAGTGGTCTGAAGCTGTCACTGGGGGCTTCACAAAATGAAT
GCCACTGAAAGCAAACGGCAGAGATCTGTATATGAAATTCTTTTTACCT
TTGTTTTCTGTCTTCCCACATGATCTATGACTGCTGCAAACCTCTATTCTGAC
10 TCCTCCTCAGTCTTGGCCGGTCTGGGAAGAACTTGGTTTAGATGCTGGGA
ATGATCCCGGGCAAAGGAAACAAATATATTTCCAGCTTGCTTTTCTTCTTC
CCCTGTGGTGCCTGCATCCTCTGTCTCGCACTGACTCACAGGCTGGACTGA
CAAAGTGGCTGTGTTGACAGGGAGGATTTAGTGGATGCCAGGCCTGACCC
TCAAAGGCCCTCAGGTGGGGCCAATTTCCAAGGTTCTTAGAGAAGATTTG
15 GGGCCACAAATCTGCTCTGCCACTTCACTGCTGGGTGAAAGCTGTTGGCTT
GAGAAGAGAGAGAATAGGTGAGGGACAATAATTCTCAAGGCAAGCTAGC
TGATATCCTTTCATTTGTGGCTTCCAAAGACAGGGAAAGGCAAGGCAGAC
CAGGGTGGAAACAGCGACATCCTGAGGGGACGACGTCCGTTTTGCTAATT
CTGTTTCCCTGGCAGCTTTGTGCTGGACAAGATTTCTCTGGTTTCAGCTGA
20 AGAGTGAACAGCTGAGGAGCTCTGTGACTCTGAGAGAGAGAGAAATCAA
TAGCCCTTGAAAGGAAAGCACTGCTTGTTCTTCCAGGCCAGGGCAGGCAG
GATAAAGAGAACAAGGCTGTACATACCAAAGCCGAATTTCCATTAAAGG
CCAGAATCCTGGGAGGCCCTTTCAGCTTAATGACCATCTTCTGATCACTAT
CACTTAGAAAATGTTTCATGTCTCCAGTTTTCTCTGCAGATAGTGCAGACAT
25 CCTGTGCACATTGATATTTTATGCACCAGATGCAATTCTCTGCAAAGGTCC
CCTCTGGCTGGTGTCCATCCGCTTGCTCTCATGACCACTGATGCGACTCTG
GGGGCAAGTCAAGCCATTCTTCTAGACTCTAGAATCTAGAGGATGATCC
CAACTCACAGGCCTCTTGGACAAGCTTGCTTTGGATTCTTCTGCCTGTAC
TACCCAGACTGCCAGTCCCTCCAGCAGGGCAATCCAACCTCCTGCCTCTAG
30 GGTGTGGGAGTCCTGGGCGGGGTTTGAATCCACAGCAGATGGCTGGCAG

GAGAGGCAGTTCCCGAGCACCCACTAGGTGCCAAGCAGGCAGGTATCAG
GTGCAGGGCTTCGGAGGTGCAGAAGCACAGCCCCTGTCCCCCTCAGGTTC
CCAGTCTGGCCTCAGGGACAGACAGGCAGGTAGGTATGGAAACGGGGTA
CTGGAAGGCATACAGCTGAGGGAGAGAGCAGCTGCACCCTGGTCAATCC
5 CCCACTCATTTCCTGGAGGTGGGGCTGAGGCACCTGGAGCACTCTGCTT
GCTTGGGTTGTCCCTTTGCCAAAGAATGCCGGGCAGAAAGTAGAGAACAA
GCAGGCCCAGGAAGCACTAGAAGTCCCTCATGGGTGGACAGAGCTTTACC
TTCGGAGCTGTCACCTCCTGTGTTCCCTCCCAACAACCCTGTGAGGCAGGCG
GGCAGGCTGAGCAGGGGACTGTGCCCCGTTGACAGATGAGGAAAGCGAG
10 GCTCAGGGGGAAGTGACTTACCTGTGGCGTCGCAGCTAGGCAGTAGTGGA
GCGGGCACCTGCCTCCAGGCCTTCGGATCCCAGCTGCCCCAGGGCGCTGC
CTCTGAGGTGCTGTACGAGGAGGTGACTCGGGCCAGCTCATGAGCAAAGG
AGCTGCCGGGCAGGAAGGACTCACTCCCCAGGGCCTGGCAGGACGGGGG
CTGTGCCCCGAGGCTGACCTGGGGGTGCCTGGCTTCCCCAAGCTTGCAGG
15 CTTTGTCTACATGAGTCTACAACAACCGATTGAACAATCCATCAGCTGCTT
GGCCACAAAATGGTTCTGACTGATCTTGTCACTGCCCTGAGTGTCTAGCTG
GGTCATCTGGGTGTTGTGGTTATACCCACCAACTGCTTTGGTGGAACGAA
ACCGCGGCCGCGGGATCATTCTGGACTCAGATTGTTCTGAAGAAGCCCA
GTTCTGGGTGGCATCAAGTGCTTGCTAGATGGGGGGCTTGCCTTGATCCG
20 GCTACACTTGGAGGTGACTTGTTCTTGACGGCTACATACAGAAAGAGAG
AAGTGGGGATGAGTTCCAAAGGCATCCTCGACTTCGGCTGTGGCCACCGG
AGGGTAGCTCCTGGCCCAACACGGACTTCTCACCTCCCGCCCTTGGCTCTC
TACTGAGCTCCCCCTGCTCCCCAATTCCTCGCCATTCCCCTCATTTCTCTG
CCCTCAGCCTGGACTGCAGTTCTTCTGGGAAGCTGCCCCAACTCCCTAGGT
25 CTGTGCTCACCAAGAGCAGATCACACTGGACTGAAATGCCAGCTGATTTG
TCTCTTCAAGAAAATTGGAAGCTCCTGGAGGTCAGGGTCCATGTCTGCTTT
TACACTCAGTGCTCTGTATGCAGGCCTGGCACTGCCCACCCTTTGACAGGT
GGTGCAATTTTGTAGAAGGAAGGAAGGGGCCAGGTGGGGTGGGCTGGG
CTGGTGGCGGGAGCTAGCTCAGCCTCTTAGATTCTCTACCCGATGGATGT
30 GACCTGGGACAGCAAGTGAGTGTGGTGAGTGAGTGACAGACGGTGCTTTGT

TCCCCTCTTGTCTCATAGCCTAGATGGCCTCTGAGCCCAGATCTGGGGCTC
AGACAACATTTGTTCAACTGAACGGTAATGGGTTTCCTTTCTGAAGGCTG
AAATCTGGGAGCTGACATTCTGGACTCCCTGAGTTCTGAAGAGCCTGGGG
ATGGAGAGACACGGAGCAGAAGATGGAAGGTAGAGTCCCAGGTGCCTAA
5 GATGGGGAATACATCTCCCCTCATTGTCATGAGAGTCCACTCTAGCTGAT
ATCTACTGTGGCCAATATCTACCGGTACTTTTTTGGGGTGGACACTGAGTC
ATGCAGCAGTCTTATGGTTTACCCAAGGTCAGGTAGGGGAGACAGTGCAG
TCAGAGCACAAGCCCAGTGTGTCTGACCCACCCAAGAATCCATGCTCGTA
TCTACAAAAATGATTTTTTCTCTTGTAATGGTGCCTAGGTTCTTTTATTATC
10 ATGGCATGTGTATGTTTTTCAACTAGGTTACAATCTGGCCTTATAAGGTTA
ACCTCCTGGAGGCCACCAGCCTTCCTGAACTTGTCTGTGCTGTCCCTGCA
ACTGGAGTGTGCCTGATGTGGCACTCCAGCCTGGACAAGTGGGACACAGA
CTCCGCTGTTATCAGGCCCAAAGATGTCTTCCATAAGACCAGAAGAGCAA
TGGTGTAGAGGTGTCATGGGCTACAATAAAGATGCTGACCTCCTGTCTGA
15 GGGCAAGCAGCCTCTTCTGGCCCTCAGACAAATGCTGAGTGTTCCTCAAGA
CTACCCTCGGCCTGGTCCAATCTCATCCCCTGGTGGCGTAAGGGTTGCTGA
ACTCATGACTTCTTGGCTAGCCTGCAACCTCCACGGAGTGGGAACTACAT
CAGGCATTTTGCTAACTGCTGTATCCTAGGCCAATAAATGTTGATCACATT
TATAGCTGCCATGGTAGGGTGGGGACCCCTGCTATCTATCTGTGGAGGCT
20 CTGGGAGCCCCCTGACACAACTTTCTGAAGCAGAGCCTCCCCAACCCCTT
TTCCATTCCCTATACCTGACAGATGGCCCAGGAACCCATTAGAAATGGAA
GGTCACTGCAGCAGTATGTGAATGTGCGTGTGGGAGAAGGGCAGGATCA
GAGCCCTGGGGGTGTGGCAGCCCCCAAGTGATTCTAATCCAGATCCTAGG
GTTGTTTCCCTGTCCCATTGAAATAGCTGCTTTAAGGGGCCTGACTCAGGG
25 AAATCAGTCTCTTGAATTAAGTGGTGATTTTGGAGTCATTTAGACCAGGCC
TTCAATTGGGATCCTGCTCTTAGAGTTGGATGAATTATTTAACTGATTTTC
AGATCTCCTCTTTCTCAATGCTTTCAGAAGCACAGTAACTGCTTACTCTGA
AATGAATTCTACCCCACTTCCACATATGCACCCCTTGCCCACCCCTTTGG
GAACACTGGCCTTAACTGCTTACCTTCAAATGGACTCATCTGTTGGGAGAT
30 ATATGCATTCTGCCGTTTCAAGGGGTCATTGCCATAAGACCTGATCTCTGTTC

CTCTTGCTAAACAGAAGATGAAAAAGACAAATTAGATTACAGCTACCAAT
TAATAATTAGCCTTAGGATCGCTGCGTGGGGACCTAGGACTTGGCTTTGG
TGCAGCAGAAAGCATGAATAAACACACCAGCATACACTCGCATGCATGCC
CCACCCTCTCGAGCAAAATTCCACAGGTATAAATAAAGTAAGATTCTGCA
5 CCTGGGTAAAAACACAACACTGCAACAGCATAGAATGGGGCAGGAGAGAC
AGAACTTAATAGCAAGAGCACACAGAAAAAAGTTTTAGGCATTTTGGATG
TCCATCTGCTCAGGATGGGTCAGCAGTGAGATGCGGTCACCAAAGAACA
AATGTAACATTAGGCTGCATTAATAGAAGCAGAGTATGTAGAAGGAGGG
AGGTGACAGTCCTATGCTAACTCTGCCTTGGCCAGACTATACCCACAGGA
10 GTCTGGGCATGCCAGTCTCAGGGAGACCCAGACAGACTGGCTGCATTCAG
AGGATGGTAAGTAATGAGAGTGGGGATTGGACTTCAAACACTACCCAGACA
AAGAATGGCTGAGCAAGCCAAGGATGCTGTGGCTGGGGCAGAGCAGACT
GTGGGCTATGTAGTGGTGGATACCTAGCCTCTGCAGGGCTGTCATAGGGA
AAGGACATTGAGAAGAGGACTGAGGCTTGTTCTTGGTGGTCCTGGCATGA
15 ACGGCCAGATGATCACATGGTCAGGTGGACACAGTCTCCAACACTGGGAG
TAGCCAAACACTTACTGCCAACCTCCCGCCCTTCTCCTGACTAGTTGCAGC
ATAGGCAATTGGGAGGAGCTTCCTGTCTCCATCTGAAAGCTGGCTGGGTG
GGCAGGGGGAGGAGCGAGCCAAGTTTCAAGGCCGCAGTTTCAGCACTCA
GTCTGGGATCGGCTCAAGGAGCAAAGGGGAAGAACATAGCCAGGAGGGA
20 ATAACATGAAGGCCCCCAGACCCAGAAAAGGCATGACTTGCTCTGAGACC
CTCAGCCGGTTGGTGTGAGGTTGTGACTCGGATCCAGGTCTGACTCCCAGT
CCAGTGCTTGAAGCCTCACCCACACAGTGAGGGGAGCCCGGCCATCTCT
GCTCAACTGCTGCCATCTCTCTCCCCTTCTCAACCACCAAGGCAGCTCTGT
CTGGGAGCACAAGCTCCAAGTCCACTTTCTGGTCTGTGTCCCCCCAAGAT
25 GCCAGAGGACTTGCCCTCTACAACACGGGCTGCCCCTGCAGTGCCTGCTTT
TCCAGCAAAGGGCTTCTGGGAACCCTTCTCTGCACTCAGTGGGGCTGGTG
GGAGTGGGGCGGGGTAGCGACCCAGTGCTTGGGACTGTGCCAGCTCTCA
GGCCTGGCAGCAGTTCCTGGCCTTGGTTCCTGCCAAGGCAGAGAGGACAA
ACACATGGCACCGGGAAGACTACACCAGAAGCGATTCCACCAGACTGGG
30 GTTTGCTTTTCTATCCCGCCCTTAGCCTGCTTCCTGTCCTGGTCCCTGCCTC

CCCCTCCACTGGAGCTGCCGTGTGGGCAGTGAGGGGCTGTTTCTCAGCTG
CCCTATGGAGCTGCCCTCTCCCTGCCAAAGCATTGGCAAGGCGGCAAGGG
GTGGGGGTGGGGATGGGGGGTGGGATCTGCCTTCTCAAGCTCTCATTATA
CTGAGCACGTCTCACCCATTATTTTATGTCATCTAGCAACACCCCATGTGG
5 AACTGAGGAGCATGGGGGTCACATGACCACTGCCCAAGGCCACACCATC
CGGATCTGCCTGAGATGGTCAGGGTTGGCAGCCATTTCTGAAGGCAGTCC
TTTCGCTTTGGCTCTTCTTGTACCAGTCTCAGGACATCAGGGCAGAAGATC
TACAGTCCCCAGCTTACTGATGTGACAGCAGAGGCTCAGAGAGGTTAAAT
GACTTGCCCAAGGTGACACGGCTAAGAAGTACAGTATCTCCTAACTGCAG
10 ACCAGGTGCTTCTGCTGCTTCTGGGGACAGATTCCTGCGTGGCTGGCTAG
GTCTAAACGGTCCTTAACCTCCATCCCCACCGGTTGCTGCATTAGTTTCATC
AAATAACACAGTTGTACAGAGGTAGGGGTTTCAGGGGCAGGGGCAGATGG
AGGCTGGAGAGTGTGACTAAGGAAACAGCAGGGGAAGTGCGGTAAAGTC
CGAAGGGAGGGACGGAAAGAGAAAAGCCAAGCCCAGGGGCGTGCCAGAC
15 AAAAGGAAAGGCCACGCCGGGGCAGGGCAGGCTTCAGCGGGTGCTGGGG
CGTCTTCATCCCGGAAGCACACATTCCAGAGGACCCCGGAGTCTAATGG
AAAAGCTGGCCAGCCTATCACTATGGAACTGCCAAGGCCACACAGCGCT
ACGCGTATTTAAATGTGCTGGAGTGTTGAGATACTGTAGTGGTGGGATGC
TTAAGTGCTGGGGTGCTGGGTGTTGGGATGCCTAGGTGCTGGGGTGCCGA
20 GATGCTGGAGTACTAGTGTGCTGGGATGCTGAAGTGCTGGGGTCCTGAGA
TGCTGCAGTGCTAGGGTGCTGAGATTCTGGGCTGCTGGAGTGTTGGGGTG
CTGGGATGCTGGAGTGCTGAGATGCTTGGACAATGGGGTGCTGGAATACT
ATGGTGCTGGGGTGCTGGGGTATTGAGATGCTAGGGTACTGGGATGCTGA
AGTGCTGAGATCCTGGAGTGCTGGGCTGCTGGGCCACAGGCTCTTGAATC
25 CATTCGTCTGCCAGGGGAAGAAACCAGAAGATAAAGAGCTAATGAAGG
AGCTTTGGTTGAGAGGGAGGAAGTAATGGAAGGAGCAACATCTTGTGGA
GGAGCAGGAGAGAATGGACCTCAGGTTGGGAGAGAGGGCCAGGCTACAG
GCCAGAGAGGCAGAAGGATTCCAGCAGAGTGTGGGCTCCAGGAGCCAAG
GGGAAACAGGTTTCTGGGAGGAGAGAGTCCAGTACTGCTGAAGTGGCAA
30 GTCCGCTGAGGACCAGGAAGCTTCATTTGGCTTTATGACCAGGAGGAATT

5
10
15
20
25
30

TGGAAGTGTGACTAGAGTACTTAGGGGGAAGGAGGCAAGACTGGAGCCA
GATTGCTCTGGGTTGAGGGGTGAGTGGGAGGTGAAGCAGGGCACTGTCAC
TCCTTTGAAGGGTGGCAGAGAGCTGGAATTGGTGCTGGATGGGCTGTGGG
GTGACAGGGTCATGTGGAAAGCCCCTGGGGGGCACCTGGAAAAGGAGAA
GCTGACAGTACAGTGAGAGGACAGCTAAGGGAAAGCGGAATGGCAGAAC
ACGCACTGCCAGGAGGAATGAGGATAGGGTCAGGAGTGCCAGGGGCAGT
GAGGCCAGCCTGGGGTCAGGTGGCAGGACGTGTCCAGGAAGCTGGTCTG
CACTGCAGCCCACACTGGCTCAGCCTTGAGGTTCCCTGTGTGGTTGGGGT
AGGAAGTTGAACCCTCTGGGAATGGAAGATGGAACCAGCTCTGCGAGCC
AAGCTCAGCTTTTATCTATGGGTCTCTGAGGGCTGGCAGAGCTGAGTGGG
GACAACTGTGATCCGTGAGGCTCTCAGGTTGAGGTGGCCCCCTCCGGGAGG
GCTTCATTTTCCCAGCGGGTAGGTTCTAAGCAGCAGTGGCTGGGCAGGTG
GGTCCAACACAGAGCCAGAGAAGGGTGAATGGGCCTCCTGGCACCCCAC
CCCTGCTGCCCCCTGAGCTCAGTGATGGAGGGGGACAGCACAGCTGAGCCC
AAGTGCTTTGGTGTGGCCCTGAGGGAAAGCTGCAGCCTGCCTGGGGCCTG
GCATGGATGGGACACTTGAGGCAGAGGGACAATAGTGGGCGCTGCAGTG
AGGCTGGCTCTTGGAGAGGTTTCCTGAGGAGTGCTGCCTGAGACGGGCAG
GGAGAACAGAGACAAAGTTGGTGACAGGGAATGAAAGCTGACTGAAGGA
CTTTACCCAGACCTATGAGGATATCTCTCTCAGCAGGAAGCAGGAGGGGA
CTGTGTGAGGACTGGCCAAGAGCTGGAGTGTTGGGAAAATGACTCTTTCT
CCGACCCCTCTGTCCTAGCTCTGGCCCCCTGGACTGCGGAGGTCTGCTTCCA
CCCCATTGGTCGATCGTTGTCCCTTGTCACAGCCATTGAGAATTTTGGCA
GGGAGCATGTTCTTAGAGCATTTTTAGGCTCTGCGGGACATAACAGCTCT
GCCTCAGAGCACATGCCTTTCTCAGCTCCTGAAAGCCACTGATCAAATTG
GAACATTTTGTACCTTAGGGATGAGGATATCAACTCTCCCAGCCACTTAG
AGGGATAAATGTGATGATGCATTCAATTGTGACTACATCTGATCCCAACT
GTTGCTTCAGCTGCTCTCCTATAGCACATGGCGGGAGGCGTGATCCCAG
TAGCTACCTCCCCACTTTTGGGGAGATGTGGTTCCATCCATGAAACCTGGG
TACCCGCCTACCAGGTCCTGGCCTATCAGGTGGCAGGGTCTGGTCAAAGA
AGGGCATGTGTGGTCTTCAGCAAGGGAGACAGGACGGTGGTGCAGAGCG

TCTAGACCCTCAGGGCAAGTCTCCCCCACACCTGCTCCCGGGGCAGTTGT
CTTTGTGACCTCCCATCCCCCTCTGTTTCATCCTCTATAAAATGAGGGGCT
GAGCCCCAAAATAACAGGCTTCTTTGCCATGATGCAAAACTGCTGAATCT
TTCTTTCTGACACACAAGGCATCGAGCAGCCTCTGAAAGAACCAAAGCCA
5 CTAGCAGGCTTCCTGACTTGGGTTTGTAGGTACTGAATACTCCCTTGAAAA
ATAAAACATAGAGGCACTTTTCTCCTGGCTGTTTATTACAGAACGAAGA
AAAAACACACTGGCTTGAAACAGACGCCAGATTTCAAATGTAGAGGTGA
AATACGAGGTGGCAATTAAAATGTGATTACAGAAAGTCTGGACACTGAGA
AAAGTTTACAGGACAGTGGGTGTGGGTTTTCTATAACAGACACTTAAATA
10 TACATGACGATAATTGCAGATAGAAACCATCAAAGACAAACCCCAAATC
AACTAATAATGTTTACAGATGTTCCCCCCCCAAACCACAGAGCCTTACATC
AAAACAAATACTGAAAGGCTTTAAACCAGGAACAGCTCGCCTTAACCCCA
CGAGGGTGCACACAAGCTGGGCTTTTTCTCTCGGTCTGAATGGTAAAGGG
AGGAGGATACTCTAGCTCCTCCAGGTGGATTGCTGAGACAGGGCTCGGCT
15 CACACACTGTCTCTGCGCCTCTCCCAAATCTGGAGAACTCTCCCAGCCTCC
TGGTAAAGTGTCTCTGTGGGGCACTTAACGATAAAACAGCTTCTGCTGTA
AAGCTCATTAGGAAAGAGCTAGCGGAGACTGAAAGGTTTCGCAAAAGAGA
TTAAGAATCACACAAGGCAATAGGATTTTTAGTGAACATAGAAATAAATG
GCCAAGTGGTTTTCTATTTGGCATTGTCAACTGCACAACAACCTCTTGGT
20 CATATCCACATTGCTCATTGCATTAACCAAGCGACTCAGCCACCTA
GCTTAACAAGGTATCACTGGAGCAAACAACACGGTCTGCATATTTGTAAC
ATTGTATAATAAACACAAAACAATGCATAGTAAACACAACCTCTACTGAAA
CAAAAGCCGTCGCTTTATTTACAAAGTCACAAAATGAAGTATAAATACTT
CTGTCATTAATGTTTAGGAAAACCATTTACAAAATTTTCAAATATGTACAC
25 GTAGCTTGAAAAATCACCAGCTTTCCATTTTGTACAGGTAGAGAGAGGG
ATAAGCATGGGCTGACAACCACTCAAATTGTAACGGGAGACAACCTGC
GGGTATGGATCGACACCACTTCCTAGAGTGATGTCACCATGGGGGTTTCT
ATGGGCATCCTGCTCAGATTTAAAGTGCCCCAGCATCCTGGGTGACTTGC
CCAGAATTCTGGGCTGTGGCATTTTGAGCAGCAGCATGCTGTTCCAAAAT
30 GTCGTCGATCAGCCTCAAGTTGCACACCCAGTCTTCATCTGGGCTCACACA

GGAGCCTTTCAAGAGAGCTTCAATGAAATCTACCTCATTGCAGTCAGGTG
ACGAAATCAGATCATTTAGTGGGGGTTGGGGCTGGCGCAAAAAGTCGGCA
GGTGGCAGCTCAGGGGGAATATCCGTTCTGTGCGAACGGACCTGGGAACTG
GCTGGCAGCAACGGCAGAAGCAGCAGCAGCGGTGGCAGCAGCAGCCACA
5 TAGCTTGGTGGCTCGATGCCCTGTATGGGGCTCAGGGGACTAAAGCTGGC
CATAACCCTGCTGGAGGAACTTGGTGGTGTGTTGCTACAGGCACCGGGCCCT
GTACCGGGCTCTGCCTGAGGCTCTGGCTGCCCAGCAGGCTGAAGCTGGGG
TTGTTGGCCAGGGGCACTTGTGTTCCCATCGCAGCGGGCACTTGTGCCTCC
CAATCAGATGGCCTCTGAAGGCAGGCCTGGCCAGAAGGTGAGTGCTGCTG
10 AACGCTATTATCCACTTGGCTGAGGGGTGTTTTCCCCGAAACTGCTGTGGT
CACAGCTGCTGCCGCTGTGACCCATGCAGCATTGTTGAACGCAGTGGGCA
TTCTTGGCACACTAGGCCGTCTGAGCTGGTGGGGACTCAAGGACTGGGTG
CCCAGGGAGCTGGGACAGAACCCAGGCAGGGGCACTTCTGGTGGGGTGG
CCTTGGGGCTCTGCATATGCTGGCAGACAGAGTCAAGTCTGCCCAGGGGA
15 GTCTGGCCTGAGTGTGAGAGGATGGGACACTGGGGGCTGGAGGTGAAAA
TTCCTTGCCGCTTCCCCAGAGTTGGTGAGATCACTCCCATGCCCGGCGCGC
GCGTCGACACGACAGGGACATACTGGGAATGCGCCCTTGCCGTGGAGGC
GGGGACCCGGCAGCGCTACGTATCCAGCATCAACCTGTATCCAGCATCAA
CCCGCCAAGTTCATAACTTGGTAGGGGTGAGGTTAGGGATCCTTAGGAG
20 CCCAGGCAGCCAGACTTTCTGGGGAGCCCATTCCTTTGTGTTGCCAAA
GTACCCCCAGCAGGTTGTGGGAATGTTGCCTGTGAAGAGAGTCTGTTGGG
GTGAGATCTTGTGTGTGTGCACAGGGTGACAGTTGTGTCCCATTTCCCGGG
AAGCTGTGATGGCAGCAGAACCTAGAGGAGCCTGAGAGAGTGTGGGAGA
GTGGGCCTCTGGAAGAGTAGAGGCTGCGGAGCCAGGTGCAGGGCTGTCT
25 GTCACCCAAAGGAAGAGGGACTGATGACTCACTGAGCGTGTGTGTCCCCT
GGTGGCAGCAGGCCCCATAGTGAACATAACCTTTTCTGTCCTGAGC
GATGCTCCCAGCAGTCCTGGGAGATGGAACGGTCCTTATTCGGCTCACAG
GAAGGACCGCCTTAACCTGGACAGACACAGCAAGGTGCTAAAGATGCCTTC
CATCAGAGGCCAGGTTGGAAGCTCTAAAGAGACTTCTCTTGCTGTTCTCTC
30 ACCCAGGGGAGGTTGTGTGTGTCCCGCTGTGGATTCTCATGTCTTTCTG

TGCCTGGTGGTCCTCTACTACATTGTGTGGTCCGTCTTGTTCTTGCGCTCTA
TGGATGTGATTGCGGACAGCGCAGGACACACATAACCATGGCCCTGAGCT
GGATGACCATCGTCGTGCCCCCTTCTTACATTTGAGGTAAGCGTTCCACGGG
AAGCCTCTTCAGCCCCTGAAGCTTGCGCTTCCCCTGACAGGATTCTGCACC
5 CCTAGAAAGGCAGCCTCTGTCCCTCGAGCTCACAGTGAGCCCACTCCAGG
AGAGGGGAGAGAAACACAGCCATCTCCGAGAGGGAGCTTCGGTGAAAGGA
GAGCATCCTTCCTTTCTCTTGGGGGCAGCACGTGGGGCTGGCAGGGAGAA
GAGTGCACCTTTTTAGCCATGGTGCCTCTGTATGGCTCCAGTTTCCACTCT
GGGGAAAGCAGAGTGGGATGTCAGATTTGTGTATTGGAGTCACGTGGAGA
10 ATTCTAGAATGGGAGCTGTTGACTCCTTAGAACAACACCCGGAGGAGTT
TGCCATAAACTGCTGGCACTGGGAACTTTTCAAGTGGATAGGCTATTGC
CGAGCTCTGAAGAGGGACATAAAAGCTCATTTCGAGCTTTCCCCAGGGAT
AGGTGGTTTCCTGCCTTTTTCTGGCGGTGCTGATGTTCCCTCTTGTGGGAG
CTCACGCGGGGGTGGGGTGGTGGGGAGGAACTGCCTAATGAAGTCTGGCT
15 TCCGCCTCTGCCCATTTCGGTGCTGGCATCAACCGGGACTATGTCTCTTT
CTTTAGATTCTGCTGGTTCACAACTGGATGGCCACAACGCCTTCTCCTGC
ATCCCGATCTTTGTCCCCCTTTGGCTCTCGTTGATCACGCTGATGGCAACC
ACATTTGGACAGAAGGGAGGAAACCACTGTATGTACTCAGCATTTAGAA
GTCCTTGGTGTGTGTCTGGGGGGGGACCAGGGGGTGGGGGGTGGCGGATA
20 GAAGTCTAGGAAGGGATGAGTCCCCGAGGGCCCCAATTTAGAAAGCTTGTG
TGGGAAAGTGAGGGCTGAGGAAATTCTGGGACCTTCTAAGGGAAGGGCA
TGCCGTAACCTCTGGTGTCTGCTGGCCTGCACCGGGACTTTTCTCGCAGTG
CACGCTGCCATTTGAGGTAGAACCAGACACGGCAGGCAACCTCTCAGAGA
TCCCGTTCCCTCCTCTGCAAAATGGGGATCAAGACAGATTCTTCCCAGGCC
25 CGGGAGGGTTTGATGGAAAATCCACATCTCCCACCCAAACCTGGGATTCA
TCCTAGGTCCCTGTTGGCCGCTCTGCCTCCCCCATATCCTTGCTGCCATCA
CCCGAGTCTTGCCTGTCTTGCTTGTCTAACACTCTATTCCCCTCCACCTGCT
TGCTGAGGCAGACACTTCCAAAACGATCTCTGCAGAGGGTGCCTTCCTGG
CAAGGCTGTGGGCTCCATGGCACGGAAGCCCAGAGCATTGCCCTTCGGAA
30 AGCCAGTGGGTTTGGGGGCAGGGCCTCACTGCAGCCCAGCAGCCCGGGCT

GTGCTTGCTGTTTGTGCCTCTGCCCCCTACCCCGCACCCGGGAGCAGGGA
GGGCTTGACCCGAGCTGACACTCCAGTAGCCTACAGAGAGGAGTAGTGG
GACTGGGAAAGTGGCTTTAAGGTGGCTCCATGAGTTCAGGCCCCCTCCTG
GCCAACCCGTGCATGACTACCGCCCTCACGGATTCCAGAGGGTGACAGAA
5 ATCTTGTTCTTGGGTGGCACTGTCATCCATGAGTTTATCCTGGCTGGAGAA
GATTAGCGGAAGACACCGTAGTCTGCGCACCACAGATATTTTGAGACTCA
CTGGAGCAGTAGTTCTCAAATTTGGGCATCCAGCAGAATCCCAAAGGGC
CAGGAAAAGGGGACCGCTGGAGCCCACCCTAGCCCGACTCAGTTTCTGGA
GGTCTGGGCTGGGGCCCGAGAATGGCATCCCTAACTAGGCCCCGTGGACG
10 CTGTCCCTGCCGGTCCGGGAACCCCACTCCAAGCACACAGAGCTAGCAT
TTGCACTTCTTCCCCATTTTGGGTACTCAAGCCCTGTTTCAGGCTTTGTGACT
CAGGAGTCTGGATAAAGTATGTTATGACATTGTAGGAGTGAACTTCTTG
TTACGGAAAGAAAGTTAACAGGAAGGTCAGTTGAGCCTCGTGTGTGAAAT
AAAAAATTCTTATTTTTCAGGGTGGTTTGGTATCCGCAAAGATTTCTGTCA
15 GTTCTGCTTGAAATCTTCCCATTTCTACGAGAATATGGAAACATTTCTTA
TGATCTCCATCACGAAGATAATGAAGAAACCGAAGAGACCCCAAGTTCCGG
AGCCCCCTAAAATCGCACCCATGTTTCGAAAGAAGGCCAGGGTGGTCATT
ACCCAGAGCCCTGGGAAGTATGTGCTCCACCTCCCAAATTAATATCGA
AATGCCAGATTAGATGCCACTTCCGGGGACAGAGCTTAAGTGGACTGGGA
20 CGCACTCTCTCCGCCTTCCTCTGCCCCCTCGTTCACCCCGCAGACCAGAAC
CAGTACTGGAGCTGGGTCTCCAGGTACGTCCATCTCATGCCTTGTTTGCAT
CCAGCGCCTATCAGCCACTCACACGACGGGACGCGGAAGTGGCAGGTG
ACGGGGGTGTGTGCCAGCAGATGCGGATGCCAGGAAGAGTGTGAGAACAA
GGGGTGGGATTACCGTCTGTCTGGGAGGGGCTCCAGGTACCCCTCTTCCC
25 CGTCAGACCCACTGGGAGATGGCTGCTTGCCAGGCCCCCAGAAGGAACAT
CTGTCTATACGGTGCTGAAATCCCAATCAAAAGTATTGTTTAGAAATGTAT
TTCTCCACAGGGCTGACCTCCTGCAGCTCGCTGAGCACTCCCAGGTCCTCA
GCACTCCCAGGTCTGTGGCTGGGGCAGTCAGTAGGAACTGTAACATATGTCT
CTGATGCACCACGTGTTTAGACACAGCACAGTCCTTTTTTCTGTTCCCTACT
30 GTGGAAGTAGTTTCTCTTTGGGCATGCTGACAGCAGTTTTTTCATAGCCTCA

CGGATGAGCCCTTTCTACGGGAGTGA CTCCATGCTTGTATACAGAGTATTT
ATACAAATGTTTTAGCATCTTCATATGCGGTGTTAACCCCTAGTTCTGTAC
AGCATATTCTGTTCAAGTATTTTTTTTACAAGCTTGTGCTGTAGGCACATGC
CTTCTGCTGCAGAAGTGGACGCCCCGTGGCACACTCCCCCCCCCCCCCGT
5 GGGGTGCCACGCCTTCATGGGACATTGCCACTTCTGCCCTGGAACCTCGTG
CAGGTACGTAGTAGCTGCTACTGCCACAACGGCAACACCAAGCAAGAGA
TGGTCCATGCTTTTCTGACGTTCTCAGAATAGTGGCTAGCTTCAAACCTGA
CAAGCGCTGCTTGAAGCCGGAACACTAGAGAATGTTGCTGAGAGCAGAA
ACGGCCACGCGGGTCACGACTATGCGTGGGAAAGTCTCAAGCTTCCTCC
10 TGCCAGCAACAAGAAGGCTTTGGAGTAGGCATGATGTTTTTCACGTGTGCG
TGCCGTTTCTCCAAGCACTGCAGGTTCCACCGTGTGTCAGAGGCTGCAAG
TTTAACATCCTCCTGCCTGAAAACAAATAGGTCCTTTGCTGAAAAGAGGG
TAAAAAAAGAGCTTTGATCTTCTCAGCCAGGAGAAGAGGGTGGTGTTC
ACGCGGGCAACTGCTCGCCGGCCTACATGGGGTTAATTCAAGTCTGCTGC
15 GAGCACGACTCCGCCCTTGGCACTGGCCTCCAGCAAGCCCTGTTCTCTTG
GGGTACAGGGGAACGGGATGGTTTAGACTTTCCTGCTCAGTGTGTAAAAA
ATGTAGCTAAAGCCACTATTTTTGCTCTCCTTAAGCTGTTCAATAAACCGG
TTCCTCATTTTACACGTGCATGATGTGTATCTTCTTTGCTGGATGGGCCAG
GAAACTGGAGTGGTCCTCTCAGCCAGCCTCAGAGGAAAGAAATCTCTAGC
20 TGGCACAGGCAGCCAGTGAGTGAGGCTGGCGGCTGCAGGGGCACAGCCT
TTAGAATGAGTCCTTCAGTGACAGGTCCCAGGGTATACGGGGTAGTGGG
AGGAAGGAGGGGACGCCTCGCAGATGCCACTGTTGGCTGGGCTACACCTT
GCCACACTTGTTACTGCTTAGGAGGCTTTCTGGAGTGTTCCTTGGGTGCTA
CGACAATCTGCAGCAGACACTGTCCTTTCACCGCTCCTGGTCCTCGTTTGC
25 TCCCCAGTGATGTCAACAGCTGAGGACTGCTCACGCTGCAACAAAAGGCT
CTGCAGTCGCTGTCTAGCTTGCCCTAGTCGTCTCTAGAGTTCTGCCTGAAC
TGAAACTCAAGTGGGGTTCAGCTCATGACTTGTGGCAATTGACCAGGAAA
TTCACCAGTTGCTGTGGCTGGAAGGATTTTCAGTCCTGTGGGTGTAAACCA
GAGGCCACAGGTGGATTCTGCCTTAGGCTCATGAGATTTCCGACTTGCTGT
30 TGAAGAAAATGCCTTGTGAAGTGACAACAGTAGCTCTGACCCAAGTCCCG

GTGCCTCGCTAGTTCCTATACGTCCCACTGGATCCTCACAGCCCCGGGAA
GCAGGTGCTACTACTCTTATCCCCGGGAGGAGACAGAGGCCGAGAGAGG
TTAAGTGACGTGCCCAAGTCACACAGCTCGGCAGCGGCCGGGTTGAGCAT
CAGCAGTCTGTTTGCAGACCCCTCACTGTCACCCCTGAGCCAGTGCGCCT
5 TGGGCCCTGCGGTCAGGATGTCTCAAGCGTGGAGGCATCACCGGTTCTGTG
GCAGTCTCTGGAAGGTCACTGAGCTCTGTGCCCAGAATCGAGTCGGGGGA
GTCTGTGCAGAGGTGGCCCTGTGTGTGGGGACAGTGTGTGACACAGACAC
TGCTTTGGATGGACACCTCTCCCGTGACCTCCTAGCATCCAATCCCAAAGG
AACAACTGTTGCAGAGATGGACCGCTGGACACAAACCCACGTGCGTTTCT
10 CTGGAGACACTGGCCAAGGAAAACAAAACATGCTCGAAGGCCAACAGCT
GCATGCCCCACCGCGATGTGACCGCAGACACCCGGGGTGTAGAAGGGTCT
CTGCCTGGTGGGGGGACACGTGCAGGCCGAGGAGAGGCAGGAAGGAGGC
TGCCTCCGACTCCCCACTGGACTGCATGGCGACGGCGTGTGGTGGGGCAG
TCAGCTAAGCCATTTGCCTAAGGGGCTGTCGGGCATCTGCGTGCTGGGGA
15 CCGACAGTGTGGGTGTGTTAGGAGGATCTGTATGGAGCACATTGCTGCCT
CTGGCTAGGACAGGGTGGAAAGGGTGGCGTGGCTACAGCCTGACCCATG
GGCACCGTCCTACCCTTTGTTCTGTGCTTCCGAGTGTCAAGTCATGTGCTGG
GGTCTGTGGGCCCATGACTCAGACGGTGAGCTCTGACCTTCCTGAGCCAG
GGCTTTGCTGTAGTTGTGCCTGGCTCAGGAGCTCTAGGACAAGGGGACCG
20 CTCCAGGTCTGCATCTACGGTGTGGCAGGGCCCCCTCGGCACTCTTGTGCAC
TAGTGTCATCTTTCCCATTTGAAATGACTGTGAGGACCAGAATGTGCACAT
GCAGATGGGCAGCTACTTGTCTGCCTTGGCCCTTTATTACACAACCTTGCTG
GGGGTGGAGATGCCACCCCCCGGCAGTCAGAGCCCCTTTATGATGTCATG
GGGCTGGTTACATGACTGCCAAGGGGTGCTGCTGGCCACACTGCACTAGC
25 AAGTTTGCCAGATGGAGGACAAGCGATCATTGAGTATGGCTCGCTGTGAA
GAAAGAAATTCGAGAGGACAGGATCATGGCTTGGAAAGGGTGCCTTTCCC
TCCCCAGTTGCAGTCAGAGACCTACCTTCACCCAGCAGATCCTTCCCCTGC
CTGGGACGACCCGGGGTCCACTGGGAGCCCTAACTTGAGGCTGCTGACAG
AAGAAATCGCTTTCCAACCTCTGGCCGAGGAAGCTTCGTTTCAAGGCCG
30 CACCCTGACGGTGACGTCCCGCCCCAGGGAGAAGATAATCTCCTCTCCCT

CCCCCTTCCACAGAACTGTGGAGACTGGTCAGCAGCAACCAGTTTTTCGT
CCATCTGGTGGGATGACAGTGGGGCTTGTAGAGTGATCAATCAAAAACCTC
TTTGAAAAGGAGATTCTCAAAAGGGACGTCGCACACAAAGTGTTTGCCAC
AACTTCGATAAAGAGCTTCTTCCGCCAGCTAAACTTGTATGGCTTCCGAA
5 AACGGCGTCAATGCACTTTCAGGACCTTCACCCGCATTTTCTCCGCAAAA
AGGCTGGTCTCCATCTTGAATAAGGTAATGAACGACAAGCCTCTGGAGGG
GTTAAGTCGGTGGGCTCTGGGGCCTGGTTCGGGTGGAAGTCCCAGGACTGC
CTCCTGGGAAGTGGGCGACCTCAGGCAGGGTGTGGGGCCATCGCTGTGGG
CCTGTGTCCCCCTCTGGGTGGAGGTGACATGAACTAAGAGTGAATGTGGG
10 GAGAGGGCTGAGGATGGTGCGGGGCCCCCTCTCGAGTGTGTAAAATATCACA
GGTGCCAAGTAGCCGTATCTGCGTGTCTCCTCCCCGGGGCCAGCCATGT
CATCTGGTGGTTGCTGTGTCCCCCTGACTCCACAGCACATTACCCTGTGAG
GTGAGCAGGCCAGGGGAGTCTGGTATTTGTACCACTGTCACCCTAGCTGG
TGTCTGGAGAGGTGCTCAAGTGGAAGCACTGAAGGGCGCCTGGCGCAGG
15 AGGTGCAGATGCTCCTGCTGCCCTTGGTAGGTGGGCCCCCTGGTGTGGAAG
AGCCAGTACCCAGGGCCTCCAACCCAGCCGGGGTGCATTCTGTTGCCAGC
TGACACTGCATGGGGGAGGCCCAGAATCTTCTTCCCTCCTGGTCTGCAACT
TCAAAGACCCTTTCCGCCGGCCATGGACACCCTAATCTGCCATTTTGAGGC
TTTTTCCAAGACGGAAAGGCCCGCCACAACCTTGGTAAACCTTGACGATGT
20 GAACGCGAGTCCCCAGCTTCCCTTTGGGGACTGGGACCTTTTCCAGAAAGG
CCTCCTGGGCCAGTAGAGTTCTCTTGCACAGGGGCGTAGATGGTTGGTAG
TTGTAGTCCATCCTTGTGACTTGCAGTTTCCTTTTGTTTATTTAAACTTTGA
CTTATAGTTAGAGTTCTACTGCCATCCTTACTTTCAAAGAGACTCCCCTCA
CCTCCTCGTGAGGATGAAGAGAAGAGTGGGTGTCAAGTCTGCACCAAGAC
25 ATCAGGAGGAGGACAAGCCAGAAGCTGCTGGATCCTGTCTGGCACCAGC
AGACACTGAGCAACAAGATCACACGTCTCCGAATGAGAATGACCAGGTC
ACACCGCAACACCGGGAACCGGGCCGGTCCCAACACCCAAATCAGGAGTG
GCTCTGCTCCACCAGCAACTCCTGTGATGGTGCCTGATTCCGCCGTGGCGA
GTGACAACAGTCCAGTGACCCAGCCGGCCGGCGAGTGGTCAGAGGGCAG
30 CCAGGCTCACGTCACTCCGGTGGCCGCTGTCCCTGGGCCTGCAGCGCTGC

CCTTCCTCTATGTCCCTGGATCTCCCACTCAGATGAATTCTTACGGGCCTG
TGGTGGCCCTTCCCACAGCGTCCCGTAGTACCCTTGCCATGGACACCACA
GGACTTCCTGCACCTGGCATGCTGCCCTTTTGCCATCTCTGGGTACCGGTG
ACCCTAGTGGCTGCTGGGGCTGCACAGCCTGCTGCCTCCATGGTCATGTTG
5 CCCCATCTCCCAGCTCTGCACCACCATTTGCCCCACAGCCACCGCACGTC
ACAGTACATGCCAGCTAGCGATGGGCCCCAGGCGTACCCAGACTACGCAG
ACCAGAGCACATAGAGGGCAGCATTTGGGCAGAATATGTGCTGGTCAATA
AATGTGTCAGAAAATGAGTAATTTTCTGACTGCACAAAAAAGTCTTCATG
GTCTCCATCCTTTTTCTTTCTTTGCAGTCCAGTACCCCCCTCCGACATCTTT
10 GGCTTG CATGATAGAGACTTCAGCCTAACACGCCTCCCTGGAAAAGTGCC
CAGGAGATACTCGAGCAGGTGCTGGGTTTTAGGAAGCCCCTCAGACATCC
CGCCTCTACTCATTA AAAAATTACCTGGAGCTGGCCGGGCGCAGTGGCTCA
TGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGTGGATCACGAGCT
CGGGAGATCGAGACTGTCCTGGCTAACACGGTGAAACCCCGTCTGTACTA
15 AAAATACAAAAAATTAGCCGGGCGTGGTGGCGGGCACCTGTAGTCCCAG
CTACTCGGGAGGCTGAGGCAGGAGAATGGCATGAACCCGGGAGGCGGAC
GTGCAGCAAGCCGAGATCGCGCCACTGCACTCCAGCCTGGGCGACAGAG
CAGGACTCCGTCTAAAAA AAAAAAAAAAAAAAAAAAATTACCTGGAGTT
GCTCGCTGAGTGCTAATTGTGCTGCTAGCATTGCCGGCCTAGGGTTCACA
20 ATGACTCTGTGTTACTGGCCATTCCTTGT CATGGCGGTGCCTTTATATAAC
AGGATCTTCACTTTACATACAGATGAGATCACGTCTGGCTTTGT CATTGGG
TTAAAAAGTGCCCTTGATAGTTGTGGAGGTGACACATACTACCGTCGACT
CGAATAGATCTGCTAGGCGCCAAGCTTGT CAGCCTTACCAGTAAAAAAGA
AAACCTATTAAAAAAACACCACTCGACACGGCACCAGCTCAATCAGTCAC
25 AGTGTA AAAAAGGGCCAAGTGCAGAGCGAGTATATATAGGACTAAAAAA
TGACGTAACGGTTAAAGTCCACAAAAAACACCCAGAAAACCGCACGCGA
ACCTACGCCCAGAAACGAAAGCCAAAAAACCCACA ACTTCCTCAAATCGT
CACTTCCGTTTTTCCCACGTTACGTA ACTTCCCATTTTAAGAAA ACTACAAT
TCCCAACACATACAAGTTACTCCGCCCTAAAACCTACGTCACCCGCCCCG

TTCCCACGCCCCGCGCCACGTCACAACTCCACCCCCTCATTATCATATTG
GCTTCAATCCAAAATAAGGTATATTATTGATGATGTTT

SEQ. ID. NO. 17

5 *Helper dependent vector C4AFO*

The sequence is shown without transgenes.

AAACATCATCAATAATATACCTTATTTTGGATTGAAGCCAATATGATAATG
AGGGGGTGGAGTTTGTGACGTGGCGCGGGGCGTGGGAACGGGGCGGGTG
10 ACGTAGTAGTGTGGCGGAAGTGTGATGTTGCAAGTGTGGCGGAACACATG
TAAGCGACGGATGTGGCAAAAGTGACGTTTTTGGTGTGCGCCGGTGTACA
CAGGAAGTGACAATTTTCGCGCGGTTTTAGGCGGATGTTGTAGTAAATTT
GGGCGTAACCGAGTAAGATTTGGCCATTTTCGCGGGAAAACCTGAATAAGA
GGAAGTGAAATCTGAATAATTTTGTGTTACTCATAGCGCGTAATATTTGTC
15 TAGGGCCGCGGGGACTTTGACCGTTTACGTGGAGACTCGCCCAGGTGTTT
TTCTCAGGTGTTTTCCGCGTTCCGGGTCAAAGTTGGCGTTTTTGATTCGGCC
GCTTGGGTCATAGACTTCTTTGAGAACCAGTTATAAGCTATGGTTTTCTCTC
CACAGAAAAAGCACTTATGGTGTCTCCCCCTTCCAGCCCACCAACATTTT
ACATCTAATTTGGGGGGGTTTTCTGGACCACTTAATACCCATCCATGGATC
20 TCATGTGAAGACTCCCCTGGCTTGAGAAATCACTGTCTTGTTGAAAATGG
GAACAAAGCTAAGTCAGATAGCTGGTTCATCAGCAATGACTTTGACCAAG
CCTGATCCCACCCTACCCACCCACCCACAGTGACCACCCCCCACAATGG
AGCACACAACCTCTAAACTGGTTTGTAGGTATGTGTGTGTGAACACGCTGA
GGAATCTGCAAAACCAAATGGTGAGTGCAAAACCAAACAGTCACGAGTA
25 AATCTCACAACAACCACGTCCTGAGCTGCAGCCCTTGTTGAACTATACCC
CACTAGGGCCCCAAGATTTTAGGACTTGTGTGTGGGTGGGACCTCCCCTTT
CTATCATGCTTTAGAAGACAGGGATTTACCAGAAATTGAACATATTGAAC
ATATGACCCATTTTTTTTCAGCCAAAGGCAATTAAAATAACTTCATACTGA
TATCCATGTCAGCAAAAGCTGCAAAACGCAAATGGGTGGCTGCTAAGAGC
30 CCTGGTACCCTGACGAGCACACCAAGTGCTTAGCAACAGTGGTGTCCAAA
GGACCAGCTGGAAGCCTGCCTTGATGAGAAGTTGCTCTTCTTTCTACATGA

AGGAACACCTCTACTCTCCTGCTTTTAATACCTGAGCTGTGAGTGATCATC
TATGTCCATTAGCAAACATCCCAGTGGAGAAGGAAACACTCATACCCGAA
ATCTAAGCTACATAGTTGGAATCACTTCAACTTATTGCAATAAACACTTAC
TAAGCACCTATTGTGGGCAAGTCTTTGCAATGGATAATAGTTCAGTAGAT
5 ATTTTGATGTAATATTTGAAATAACAATAAAAAATTGCCACCACTGAATTTA
TTGAGCATTGTGCTGTGCTTTAGGCACTAACCCAGGTTCTTTAAATATTTGG
TCTTATTCGATCTGTATAAATAGCCATCTATGAGAAAGGGACTATTATTGC
CCTTATTTTACAAATGAGGCCAATGAGGCCAGAGAGGTTAACTAAGTTG
CCCAAATCATACAGCCCACTAGTGGCAGAGCAGGATGCAAACCCAGGC
10 TTGCCTCGTTCCCAAGCCACATGTCGTTTGCATTGGTGTGGAGGTGTGCA
TGTGTTTAGTCATTAGCATGTTATATGATAAGCAAGTTTTGAAACATAGAA
ACTTAAAATGTGCCATTAAGAAAAGTACAGGCAAGGTTTTCCAAGGGGAG
GTGTGGACCTCCGGACAAATTTTTAAGAACTAATTATAAATACTTAAAAA
TGGGAATAAGAAGACAACCTAACTACCTGAACAGTTTTAGAGATGACTCA
15 TGCCCAACCCTCTAAAACCCAAACAAAAACAACAAAGTCAAGAAAACCCA
TGAAATCTTAGCAAGCGATTTCTATGTACTTGTGAAAAGGATTTCTTTACC
ATTCTAATGGGATTTATGCCAACCATAGAGGGCTCAGTGCCCCTCCCATG
GGGTGGTTAGTGAGTACAGAGCTGAGCTCACCGGCCATCTGCAGCTTCAT
GTTATCAAGCTCCAGTTTGTCTTGGAGCAAGGTTATCTGGGACATGAGC
20 AGAGGCATTGCTTTCTGCAATGGACAGTTCTTTCTGCCTGCATACCTAGCT
CCTTGATAACTTTAAATACCATTTTATAGCCACACTGGAGTTTTGAAGACC
TCAATATGCAAATATTACTCAGGTTCTGATTACTTGTCTGCTCCATGATAA
CACACTCTAAAAGCAATGAATGGGGCTTATTTGTAGAGAACTGAAGCATT
TTAAGCTTTTGCTCAGGAATCCCTGGTAGCTTCCTGTGACTTGCAAGATAT
25 TAGTGATGGGTCAAGAAACAGGACCCCCCATCAGCATAACATACGCAGTG
CCTCAGTAGTCCATCAGGCAGAAAAAACTGCAGATGGCACATGGAAATG
ACCAGCGGCGGAAGATACCCCGACAGTGTGGGCAGTTCTATTTACGCAGC
AATCAAGAGGGGGCCTGGAGCCACTCAATCAAGTGGAGCAGGATGGGAG
CAAGCACTGTGCAGACCAATGCAATCCCCAGTTAACACAAAAAATAAATA
30 AAAGAGATGAGATTCAGTCTCTTGACTGTGACTGACTGGGAGCTTTATAG

CTGATGCTTGTGTCTTTTCTCCATTTTATTTAATTAGGAAAAGAAATGCTT
ATCACACACTCTACGTGTGAGGTACAACCTCCACAGGAAAGGTTGTTAGG
AACATTTCAACTTCTAGAAGTTTCTAAACATAAGGTAAATCCATCTTTGTC
CTTGGGATCACTGCACATCTCAGAAAGGCAAATAAATCAGTAATTGGTGG
5 GCATAATTACTAGCTCATGGACTGACAAGGTCTACACTATTTCTGAATCTCA
CAGAAGTAAGCCATGGGACAGATAGAGTCTGATAGTGGTGCCCCGTTTCC
TGGAGGTCACACTTACTCATCCCCCTGGACCCTGGGCTTCTCATGATTGTC
AGAGAGTTTGCTGGAACCAGGTGAGCCAGTTTCCCTTCCCCTGAAAAAT
CCTCCAATGGCTCGCACCAAGACTAGAGATGCAAGTGCAAGCACATCCAC
10 CCTCTCAGCAGCCAGGTTTTGCGTTCCATAATGTCACGTACCCCCAGTCAT
ACCAATCTCCTTGGAGCTCTCCAGACAGGCTGCCATGTGTGGTCGGCCCTC
TGTGCTTGTGCTCCTTGGTTTGCCAAGCCTGGAATGCCCTTTCTCCATGAT
TGTACTCTGGGAATTCTGTGTGTCTTTCGAGATGAAGCTCCTCCACCCTGG
AAATTCTGCCTCTCTTTCCAGGTCCAGCTCCTGTGGCTGAATACCCTTGGC
15 TGACTGAACATTCTTGCTGGGCCAAGTCTTAATCATCTCTAAATCCCTCTG
GTGCCCCGACAGTATCTAGCCCAGGCAAAGGGCTCAGCAAATACTTGCAA
AATTCAATAAACTTTACTATGTTACTAGCTTGGCATGATGTTCTAGGCACT
TGGGAGATTATAATCTGGTGAGCATTGTGCTAGTCTTTTTTTTGGGATAAA
CACCAATGATAAACTGAGCCTTTACTCTCATTTTAGCACATCTACTCATCT
20 CTTTTCAATGCTGGTGGGTTTATAAAACCCATCAGTAGAACAGAGAGTGA
TCATAATCATATGGTGAAAATAACATCAGCTAACATATTTACTGAACATG
CTTTAGTGTGCTGGGCACTGAACTCTGTGCACATGTGAATTTGAGATAGAT
TGTTCCCTAGCTAATAAGATGAGGAAGTGGAGATGGGCTTACTCAAGTCAC
ACAATAGCAAGATGGGGTACAGACAGAAACCAAGCTAGAAACCAAGCAA
25 CCCCTGGGTTTGGAATGCATGGGCTCCTCCTCTGCACATGGCGAGGAGC
AGTCAGGTGCTCCTCCTTCTCTCATACTGAAATAACTCTGCACTTTTGGCT
ATTCTGTGACACTCTGTGCTATTCTGTAGCTCAAATGGCTCTGGTGCAAGG
AGCTCAGATTATGACTAACCCCCATAGATATTCAGCTGCTTTGCAAGAAG
TGGATGAATGCTCTGGCTTTATAAATTATTGACTAGATTGGATATTGGCCC
30 AATCTCCACTCTGGTGATTTCGGAAGGAGGCATATGCACATTTGCAAGGTT

ATAGTAGTGCCTCTAATTCCACTGGCTTCTGAGAGCTTGTAGGTTTCTTG
ACTTAATCATTCTGGAATTAAGGTAATGGATCCTCAACACCTTTTCTTTCC
CTGGCCTTTACTAACCATGTAACAGAAATGAGCAGAGAAAACCCAAGAA
AGCGAACTGGAGACTTGATGAGTGTGTCAAAGATGCTCAGAGTCCAAGGT
5 CCTCTGTGGCTCACTGACTTCAGAAGCCAACCTCCGTTGTTCAAGTCACTT
GTGAGGTTACTATGCTAGAGCACTAAATATTATCCAGATAGCCCAAAGAG
GTGAAGGCAGACATGTGGAAGAACCTGGATTTTTGGGCAAGTTGATGAAT
GCCTCCTGCCCCATATCAAAGAGAGGTGATAGGAGACAACCTTTGTGAATT
TGAAATAATGCACCGGGGAAACAGGAAAACGTAATGTAAGTCACGCTTCT
10 TGGCTTTTTTCTTGCCATTACCCTACTTGGCCAAGTGCAAATGGGATTTC
ATATATCATAAGTATGCATCTATTAATAACAATGCAAGAAAGCTGTACAA
CTGAAGTCTGAGATTTTGTAAAGAAACAGAATCTCTGTAAGCATCACCATC
CAACAGAACTTCCTGAGTTGATGCTGAATCATCCCAGAAAGAAGGCGCGC
CAGCTCTGCAGGATCTTCAAGTCTGGGGTGCCACCAGCAAGCGACGGTCC
15 TCCATGGGCTCTTCACCTTACGGCAGTGTCCAGAGGCACCGCCAGTCTCT
GCTCCTATGCTGGTCCTGCTGTCCCTGGCAAAAGGAGCCAGAGCATTCTCT
CCAGGCCTCCCGAGGAGGCTGCTTCCTTTGTTTTGCAGATGGAGGCTCCCA
TCCTTTGTTCTGAATCAATGTGCTCCAAAGATAAGCCCCAAGAAAACAGT
TGTTGCCTTTTGACACTGACAATTAGAATCGTTGGAAAATGGAGAAAACA
20 GGAAATGGCAAATGGTTTCAGTGACCAGGAGGAAACCGTGCCTGAAAGTT
GCTGCTTAGTGACTGGGACACTCGCTTTCTGCTCTCTTATGAAGGACAGCC
TAGGCCGTGTGGCCTTTTATAAACAAAGCTATGAAGGGGTCGTCAAATTT
TCTAGGGCTGCAACTGTGGCACTACGTCCTGTTGTGCCAGGTGACACTGA
CAAGCAGCACTGAGTTCTATGCAAGCCCAGGTGTGCTTCTCTCATGGTGA
25 CCCCCAGAGAACTAAGGCCCAGCTCTTCCTCTGTACACCCCTCCCAGCC
CCCACTGTCAGACAAGGGACCACATTCACAGACAGTCTCAGCCAAGATGG
CAACCTTGGAAGTCCTGGGGATGCCTTTCTAGAAGCTTTAGGCTGACGCC
AGGGAGCTGTAAAGCCCCCACCTGTCCCTGGGGGTTGTGTGCTGGGCAGG
AGAGAGGGGAAAGAGCCCAAACCTCCACCACTGCCTGCTGGCACAACCTGA
30 GCCCACGCCAGGCACTGCACCAGCTTCACTCACCAGCTACCACATGCTAA

TGTCTCCTGTTTACCCAGGTGCAATCACCTGAGTGCTGCTTTCTCACCCT
GCATGGGAAGAACGATCTCATTACCAATTCAACCCTTAAGCAGGCAAAA
GGACTCCAACGGGGGAAAGGCAGAGGACAAACGATTCGGGGACTGAGAC
AGATGCCCACAAAATGTTCTGGATTTTAGCCGGATTCGGGGCTGAAGTTC
5 CTTGCCCTATGAGCCCTTGGAGGAGAGCATCCTGACTCAGACAAGAATTC
AAATGACAAGACTTCTGGGGGATGGCAGTGAATGGATCCAAAGCCTTGTT
TTATAAAGAATGTGGAGGAAGGGAGGAAGGAAGGGAGGGGAAGGGGAGA
CTGAGAGGGAGAGAGAATAAAGAAGAGGAAGAGAAGGAAGAGGAGGAG
GAGGAAGGGGAGGGGAGAGGAAAAAGGGGAGGGTAGAGGAGGAGGGAG
10 AAGGGGAAGGGAGGACAGAAGAGGGGAGGGAGCGGGGAGATTATCTAC
CCCTCCATGCTAGGGAACTCCTGTCTCCTTTTCCAGCAGACACTGCTCCAT
GAGTGCCCCCACAGAGAGCCAGCCTGCCACACAGCAGAAGGTGCTCAGC
TATCTGTCCCAAAGGTTGAGCACTAGTGTATACCACAAGGAAAAAACCT
TCTTTTGTACTGTTTGGAAGCCAGAATTAGAAACAACCTCGTTCACCCTTCT
15 GACTTTCCCTTTGCAATGACAGGAAGATAGTATACTGCATGGGACCACCA
CACATTCAAGATCTGAGTTGTGGGGACCCCCACCCCCACATGGGTGGAGC
CCCCTGAGACCCGTGGGCTCTGACTGGCCATGGCCCATAGATAGTGACAA
CATAACACTGCTATCTGAAGCCTAGTTGGAGGACGGATGGATGCTTTTA
GGAAAAAATTCCCTTTAGTCGTCCCTCAGAACTACAACGGCTTCCTAAA
20 ATGAATGGCAGGCAAGTTTTGATTGTTTCCTAAGTTCCCCCTAACTCTCAG
AGCCTGTGACTCAGCTTCTCCAGAATGCCTGCAGGGAGGCAGGACCACTG
GGAGTAATGAGCACGAATTGGGCCACTTCTCCCAGTGGACTGTTAACGCC
GAGATCTGGGGTGGAAACCACTGAACACCTGCCCTTTTGCGTTGCCCAGA
ATGCAGCCAGAGGGATACGGGCCACAGTAACTCTGGGGCCCCCTGAACCTA
25 CCTTCCCAAGACCCTGACGGGAAGGTATTGCTTACATGACAGTCCAGTCA
GTCCCTCTGTCACTCAGTTAACAACATGGTCATGGGTCTACTATGGGCCA
GGCTCTGTTCCGGGCATGGGGATAACACCGTGTGTCTCTACTCTTATGGAC
ACAAGTAAATCAATCAAAGCAGATAATGTTCTAATGACCACCCAGTAGGG
TGAAGGGATGGAGGGTGGGCAGGTGCTGGCAGAAACCTAAGGTCAAGTC
30 ATGGGACCCCAATGGGGAACTGGAAACTGCCACACCATTTTCTATGTTAG

GGGTCTTTCCTTGCCCTTCTAGAGACCTCACCCATGGTAATGATGCTGCCC
TCAGGTTTGAAGCTGGTCTGAAGCTGTCAGTGGGGGCTTCACAAAATGAAT
GCCACTGAAAGCAAACGGCAGAGATCTGTATATGAAATTCTTTTTACCCT
TTGTTTTCTGTCTTCCCACATGATCTATGACTGCTGCAAACCTCTATTCTGAC
5 TCCTCCTCAGTCTTGGCCGGTCTGGGAAGAACTTGGTTTAGATGCTGGGA
ATGATCCCGGGCAAAGGAAACAAATATATTTCCAGCTTGCTTTTCTTCTTC
CCCTGTGGTGCCTGCATCCTCTGTCTCGCACTGACTCACAGGCTGGACTGA
CAAAGTGGCTGTGTTGACAGGGAGGATTTAGTGGATGCCAGGCCTGACCC
TCAAAGGCCCTCAGGTGGGGCCAATTTCCAAGGTTCTTAGAGAAGATTTG
10 GGGCCACAAATCTGCTCTGCCACTTCACTGCTGGGTGAAAGCTGTTGGCTT
GAGAAGAGAGAGAATAGGTGAGGGACAATAATTCTCAAGGCAAGCTAGC
TGATATCCTTTTCAATTTGTGGCTTCCAAAGACAGGGAAAGGCAAGGCAGAC
CAGGGTGGAAACAGCGACATCCTGAGGGGACGACGTCGTTTGTCTAATT
CTGTTTCCCTGGCAGCTTTGTGCTGGACAAGATTTCTCTGGTTTCAGCTGA
15 AGAGTGAACAGCTGAGGAGCTCTGTGACTCTGAGAGAGAGAGAAATCAA
TAGCCCTTGAAAGGAAAGCACTGCTTGTTCTTCCAGGCCAGGGCAGGCAG
GATAAAGAGAACAAGGCTGTACATACCAAAGCCGAATTTCCATTAAAGG
CCAGAATCCTGGGAGGCCCTTTCAGCTTAATGACCATCTTCTGATCACTAT
CACTTAGAAAATGTTTCATGTCTCCAGTTTTCTCTGCAGATAGTGCAGACAT
20 CCTGTGCACATTGATATTTTATGCACCAGATGCAATTCTCTGCAAAGGTCC
CCTCTGGCTGGTGTCCATCCGCTTGCTCTCATGACCACTGATGCGACTCTG
GGGGCAAGTCAAGCCATTCTTCTAGACTCTAGAATCTAGAGGATGATCC
CAACTCACAGGCCTCTTGGACAAGCTTGCTTTGGATTCTTCTGCCTGTAC
TACCCAGACTGCCAGTCCTCCAGCAGGGCAATCCAACCTCCTGCCTCTAG
25 GGTGTGGGAGTCCTGGGCGGGGTTTGGAAATCCACAGCAGATGGCTGGCAG
GAGAGGCAGTTCCCGAGCACCCACTAGGTGCCAAGCAGGCAGGTATCAG
GTGCAGGGCTTCGGAGGTGCAGAAGCACAGCCCCTGTCCCCCTCAGGTTT
CCAGTCTGGCCTCAGGGACAGACAGGCAGGTAGGTATGGAAACGGGGTA
CTGGAAGGCATACAGCTGAGGGAGAGAGCAGCTGCACCCTGGTCAATCC
30 CCCACTCATTTCCCTGGAGGTGGGGCTGAGGCACCTGGAGCACTCTGCTT

GCTTGGGTTGTCCCTTTGCCAAAGAATGCCGGGCAGAAAGTAGAGAACAA
GCAGGCCCAGGAAGCACTAGAAAGTCCCTCATGGGTGGACAGAGCTTTACC
TTCGGAGCTGTCACCTCCTGTGTTCTCTCCCAACAACCCTGTGAGGCAGGCG
GGCAGGCTGAGCAGGGGACTGTGCCCCGTTGACAGATGAGGAAAGCGAG
5 GCTCAGGGGGAAGTGACTTACCTGTGGCGTCGCAGCTAGGCAGTAGTGGA
GCGGGCACCTGCCTCCAGGCCTTCGGATCCCAGCTGCCCCAGGGCGCTGC
CTCTGAGGTGCTGTACGAGGAGGTGACTCGGGCCAGCTCATGAGCAAAGG
AGCTGCCGGGCAGGAAGGACTCACTCCCCAGGGCCTGGCAGGACGGGGG
CTGTCGCCCCGAGGCTGACCTGGGGGTGCCTGGCTTCCCCAAGCTTGCAGG
10 CTTTGTCTACATGAGTCTACAACAACCGATTGAACAATCCATCAGCTGCTT
GGCCACAAAATGGTTCTGACTGATCTTGTCACTGCCCTGAGTGTCTAGCTG
GGTCATCTGGGTGTTGTGGTTATACCCACCAACTGCTTTGGTGGAACGAA
ACCGCGGCCGATCTGAATTCAGATCGGCCGCGGGATCATTCTGGACTCA
GATTGTTCTGAAGAAGCCCAGTTCTGGGTGGCATCAAGTGCTTGCTAGAT
15 GGGGGGCTTGCCTTGATCCGGCTACACTTGGAGGTGACTTGTTCTTGGAC
GGCTACATACAGAAAGAGAGAAGTGGGGATGAGTTCCAAAGGCATCCTC
GACTTCGGCTGTGGCCACCGGAGGGTAGCTCCTGGCCCAACACGGACTTC
TCACCTCCCGCCCTTGGCTCTCTACTGAGCTCCCCCTGCTCCCCAATTCC
TCGCCATTCCCCTCATTCTCTGCCCTCAGCCTGGACTGCAGTTCTTCTGG
20 GAAGCTGCCCCAACTCCCTAGGTCTGTGCTCACCAAGAGCAGATCACACT
GGACTGAAATGCCAGCTGATTTGTCTCTTCAAGAAAATTGGAAGCTCCTG
GAGGTCAGGGTCCATGTCTGCTTTTACACTCAGTGCTCTGTATGCAGGCCT
GGCACTGCCCACCCTTTGACAGGTGGTGCATATTTTGTAGAAGGAAGGAA
GGGGCCAGGTGGGGTGGGCTGGGCTGGTGGCGGGAGCTAGCTCAGCCTCT
25 TAGATTCTCTACCCGATGGATGTGACCTGGGACAGCAAGTGAGTGTGGTG
AGTGAGTGCAGACGGTGCTTTGTTCCCCTCTTGTCTCATAGCCTAGATGGC
CTCTGAGCCCAGATCTGGGGCTCAGACAACATTTGTTCAACTGAACGGTA
ATGGGTTTCCTTTCTGAAGGCTGAAATCTGGGAGCTGACATTCTGGACTCC
CTGAGTTCTGAAGAGCCTGGGGATGGAGAGACACGGAGCAGAAGATGGA
30 AGGTAGAGTCCCAGGTGCCTAAGATGGGGAATACATCTCCCCTCATTGTC

ATGAGAGTCCACTCTAGCTGATATCTACTGTGGCCAATATCTACCGGTACT
TTTTTGGGGTGGACACTGAGTCATGCAGCAGTCTTATGGTTTACCCAAGGT
CAGGTAGGGGAGACAGTGCAGTCAGAGCACAAGCCCAGTGTGTCTGACC
CACCCAAGAATCCATGCTCGTATCTACAAAAATGATTTTTTCTCTTGTAAT
5 GGTGCCTAGGTTCTTTTATTATCATGGCATGTGTATGTTTTTCAACTAGGTT
ACAATCTGGCCTTATAAGGTTAACCTCCTGGAGGCCACCAGCCTTCCTGA
AACTTGTCTGTGCTGTCCCTGCAACTGGAGTGTGCCTGATGTGGCACTCCA
GCCTGGACAAGTGGGACACAGACTCCGCTGTTATCAGGCCCAAAGATGTC
TTCCATAAGACCAGAAGAGCAATGGTGTAGAGGTGTCATGGGCTACAATA
10 AAGATGCTGACCTCCTGTCTGAGGGCAAGCAGCCTCTTCTGGCCCTCAGA
CAAATGCTGAGTGTTCCCAAGACTACCCTCGGCCTGGTCCAATCTCATCCC
ACTGGTGCGTAAGGGTTGCTGAACTCATGACTTCTTGGCTAGCCTGCAAC
CTCCACGGAGTGGGAACACTACATCAGGCATTTTGCTAACTGCTGTATCCTA
GGCCAATAAATGTTGATCACATTTATAGCTGCCATGGTAGGGTGGGGACC
15 CCTGCTATCTATCTGTGGAGGCTCTGGGAGCCCCTGACACAAACTTTCTGA
AGCAGAGCCTCCCCAACCCCTTTTCCATTCCCTATACCTGACAGATGGCCC
AGGAACCCATTAGAAATGGAAGGTCAGTGCAGCAGTATGTGAATGTGCGT
GTGGGAGAAGGGCAGGATCAGAGCCCTGGGGGTGTGGCAGCCCCCAAGT
GATTCTAATCCAGATCCTAGGGTTGTTTCCCTGTCCCATTGAAATAGCTGC
20 TTTAAGGGGCCTGACTCAGGGAAATCAGTCTCTTGAATTAAGTGGTGATTT
TGGAGTCATTTAGACCAGGCCTTCAATTGGGATCCTGCTCTTAGAGTTGGA
TGAATTATTTAACTGATTTTCAGATCTCCTCTTTCTCAATGCTTTCAGAAGC
ACAGTAACTGCTTACTCTGAAATGAATTCTCACCCCACTTCCACATATGCA
CCCCTTGCCCAACCCTTTGGGAACACTGGCCTTAACTGCTTACCTTCAAAT
25 GGACTCATCTGTTGGGAGATATATGCATTCTGCCGTTTCAGGGGTCAATTGCC
ATAAGACCTGATCTCTGTTCCCTCTTGCTAAACAGAAGATGAAAAAGACAA
ATTAGATTACAGCTACCAATTAATAATTAGCCTTAGGATCGCTGCGTGGG
GACCTAGGACTTGGCTTTGGTGCAGCAGAAAGCATGAATAAACACACCAG
CATACACTCGCATGCATGCCCCACCCTCTCGAGCAAAATTCCACAGGTAT
30 AAATAAAGTAAGATTCTGCACCTGGGTAAAAACACAACACTGCAACAGCAT

AGAATGGGGCAGGAGAGACAGAACTTAATAGCAAGAGCACACAGAAAAA
AGTTTTAGGCATTTTGGATGTCCATCTGCTCAGGATGGGTCAGCAGTGAG
ATGCGGTACCAAAAAGAACAAATGTAACATTAGGCTGCATTAATAGAAGC
AGAGTATGTAGAAGGAGGGAGGTGACAGTCCTATGCTAACTCTGCCTTGG
5 CCAGACTATACCCACAGGAGTCTGGGCATGCCAGTCTCAGGGAGACCCAG
ACAGACTGGCTGCATTTCAGAGGATGGTAAGTAATGAGAGTGGGGATTGG
ACTTCAAACCTACCCAGACAAAGAATGGCTGAGCAAGCCAAGGATGCTGT
GGCTGGGGCAGAGCAGACTGTGGGCTATGTAGTGGTGGATACCTAGCCTC
TGCAGGGCTGTCATAGGGAAAGGACATTGAGAAGAGGACTGAGGCTTGTT
10 CCTGGTGGTCCTGGCATGAACGGCCAGATGATCACATGGTCAGGTGGACA
CAGTCTCCAACACTGGGAGTAGCCAAACACTTACTGCCAACCTCCCGCCC
TTCTCCTGACTAGTTGCAGCATAGGCAATTGGGAGGAGCTTCCTGTCTCCA
TCTGAAAGCTGGCTGGGTGGGCAGGGGGAGGAGCGAGCCAAGTTTCAAG
GCCGCAGTTTCAGCACTCAGTCTGGGATCGGCTCAAGGAGCAAAGGGGA
15 AGAACATAGCCAGGAGGGAATAACATGAAGGCCCCCAGACCCAGAAAAG
GCATGACTTGCTCTGAGACCCTCAGCCGGTTGGTGTGTCAGGTTGTGACTCG
GATCCAGGTCTGACTCCCAGTCCAGTGCTTGAAGCCTCACCCACACAGT
GAGGGGAGCCCGGCCATCTCTGCTCAACTGCTGCCATCTCTCTCCCTTCT
CAACCACCAAGGCAGCTCTGTCTGGGAGCACAAGCTCCAAGTCCACTTTC
20 TGGTCTGTGTCCCCCCCCAAGATGCCAGAGGACTTGCCTCTACAACACGGG
CTGCCCCGTGCAGTGCCTGCTTTTCCAGCAAAGGGCTTCTGGGAACCCTTCT
CTGCACTCAGTGGGGCTGGTGGGAGTGGGGCGGGGTAGCGACCCAGTGCT
TGGGACTGTGCCCAGCTCTCAGGCCTGGCAGCAGTTCCTGGCCTTGGTTCC
TGCCAAGGCAGAGAGGACAAACACATGGCACCGGGAAGACTACACCAGA
25 AGCGATTCCACCAGACTGGGGTTTGCTTTTCTATCCCGCCCTTAGCCTGCT
TCCTGTCCTGGTCCCTGCCTCCCCCTCCACTGGAGCTGCCGTGTGGGCAGT
GAGGGGCTGTTTCTCAGCTGCCCTATGGAGCTGCCCTCTCCCTGCCAAAGC
ATTGGCAAGGCGGCAAGGGGTGGGGGTGGGGATGGGGGGTGGGATCTGC
CTTCTCAAGCTCTCATTATACTGAGCACGTCTCACCCATTATTTTATGTCAT
30 CTAGCAACACCCCATGTGGACACTGAGGAGCATGGGGGTCACATGACCAC

TGCCCCAAGGCCACACCATCCGGATCTGCCTGAGATGGTCAGGGTTGGCAG
CCATTTCTGAAGGCAGTCCTTTTCGCTTTGGCTCTTCTTGTACCAGTCTCAG
GACATCAGGGCAGAAGATCTACAGTCCCCAGCTTACTGATGTGACAGCAG
AGGCTCAGAGAGGTTAAATGACTTGCCCCAAGGTGACACGGCTAAGAAGT
5 ACAGTATCTCCTAACTGCAGACCAGGTGCTTCTGCTGCTTCTGGGGACAG
ATTCTGCGTGGCTGGCTAGGTCTAAACGGTCCTTAACTCCATCCCCACCG
GTTGCTGCATTAGTTTCATCAAATAACACAGTTGTACAGAGGTAGGGGTT
CAGGGGCAGGGGCAGATGGAGGCTGGAGAGTGTGACTAAGGAAACAGCA
GGGGAAGTGCGGTAAAGTCCGAAGGGAGGGACGGAAAGAGAAAGCCAA
10 GCCCAGGGGCGTGCCAGACAAAAGGAAAGGCCACGCCGGGGCAGGGCAG
GCTTCAGCGGGTGCTGGGGCGTCTTCATCCCGGAAGCACACATTCCAGA
GGACCCCGGAGTCTAATGGAAAAGCTGGCCAGCCTATCACTATGGAACT
GCCAAGGCCACACAGCGCTACGCGTATTTAAATGTGCTGGAGTGTTGAGA
TACTGTAGTGGTGGGATGCTTAAGTGCTGGGGTGCTGGGTGTTGGGATGC
15 CTAGGTGCTGGGGTGCCGAGATGCTGGAGTACTAGTGTGCTGGGATGCTG
AAGTGCTGGGGTCCTGAGATGCTGCAGTGCTAGGGTGCTGAGATTCTGGG
CTGCTGGAGTGTTGGGGTGCTGGGATGCTGGAGTGCTGAGATGCTTGGAC
AATGGGGTGCTGGAATACTATGGTGCTGGGGTGCTGGGGTATTGAGATGC
TAGGGTACTGGGATGCTGAAGTGCTGAGATCCTGGAGTGCTGGGCTGCTG
20 GGCCACAGGCTCTTGAATCCATTCGTCTGCCCAGGGGAAGAAACCAGAAG
ATAAAGAGCTAATGAAGGAGCTTTGGTTGAGAGGGAGGAAGTAATGGAA
GGAGCAACATCTTGTGGAGGAGCAGGAGAGAATGGACCTCAGGTTGGGA
GAGAGGGCCAGGCTACAGGCCAGAGAGGCAGAAGGATTCCAGCAGAGTG
TGGGCTCCAGGAGCCAAGGGGAAACAGGTTTCTGGGAGGAGAGAGTCCA
25 GTACTGCTGAAGTGGCAAGTCCGCTGAGGACCAGGAAGCTTCATTTGGCT
TTATGACCAGGAGGAATTTGGAAGTGTGACTAGAGTACTTAGGGGGAAGG
AGGCAAGACTGGAGCCAGATTGCTCTGGGTTGAGGGGTGAGTGGGAGGT
GAAGCAGGGCACTGTCACTCCTTTGAAGGGTGGCAGAGAGCTGGAATTGG
TGCTGGATGGGCTGTGGGGTGACAGGGTCATGTGGAAAGCCCCTGGGGGG
30 CACCTGGAAAAGGAGAAGCTGACAGTACAGTGAGAGGACAGCTAAGGGA

AAGCGGAATGGCAGAACACGCACTGCCAGGAGGAATGAGGATAGGGTCA
GGAGTGCCAGGGGCAGTGAGGCCAGCCTGGGGTCAGGTGGCAGGACGTG
TCCAGGAAGCTGGTCTGCACTGCAGCCCACACTGGCTCAGCCTTGAGGTT
CCCTGTGTGGTTGGGGTAGGAAGTTGAACCCTCTGGGAATGGAAGATGGA
5 ACCAGCTCTGCGAGCCAAGCTCAGCTTTTATCTATGGGTCTCTGAGGGCTG
GCAGAGCTGAGTGGGGACAACACTGTGATCCGTGAGGCTCTCAGGTTGAGGT
GGCCCCTCCGGGAGGGCTTCATTTTCCCAGCGGGTAGGTTCTAAGCAGCA
GTGGCTGGGCAGGTGGGTCCAACACAGAGCCAGAGAAGGGTGAATGGGC
CTCCTGGCACCCACCCCTGCTGCCCCTGAGCTCAGTGATGGAGGGGGAC
10 AGCACAGCTGAGCCCAAGTGCTTTGGTGTGGCCCTGAGGGAAAGCTGCAG
CCTGCCTGGGGCCTGGCATGGATGGGACACTTGAGGCAGAGGGACAATA
GTGGGCGCTGCAGTGAGGCTGGCTCTTGGAGAGGTTTCCTGAGGAGTGCT
GCCTGAGACGGGCAGGGAGAACAGAGACAAAGTTGGTGACAGGGAATGA
AAGCTGACTGAAGGACTTTACCCAGACCTATGAGGATATCTCTCTCAGCA
15 GGAAGCAGGAGGGGACTGTGTGAGGACTGGCCAAGAGCTGGAGTGTTGG
GAAAATGACTCTTTCTCCGACCCCTCTGTCTAGCTCTGGCCCCTGGACTG
CGGAGGTCTGCTTCCACCCCCATTGGTCGATCGTTGTCCCTTGTACAGCC
ATTGAGAATTTTGGCAGGGAGCATGTTCTTAGAGCATTTTATAGGCTCTGCG
GGACATAACAGCTCTGCCTCAGAGCACATGCCTTTCTCAGCTCCTGAAAG
20 CCACTGATCAAATTGGAACATTTTGTACCTTAGGGATGAGGATATCAACT
CTCCCAGCCACTTAGAGGGATAAATGTGATGATGCATTCAATTGTGACTA
CATCTGATCCCAACTGTTGCTTCAGCTGCTCTCCTATAGCACATGGCGGGA
GGCGTGATCCCAGTAGCTACCTCCCCACTTTTGGGGAGATGTGGTTCCAT
CCATGAAACCTGGGTACCCGCCTACCAGGTCCTGGCCTATCAGGTGGCAG
25 GGTCTGGTCAAAGAAGGGCATGTGTGGTCTTCAGCAAGGGAGACAGGAC
GGTGGTGCAGAGCGTCTAGACCCTCAGGGCAAGTCTCCCCCACACCTGCT
CCCGGGGCAGTTGTCTTTGTGACCTCCCATCCCCCTCTGTTTCATCCTCTAT
AAAATGAGGGGCTGAGCCCCAAAATAACAGGCTTCTTTGCCATGATGCAA
AACTGCTGAATCTTTCTTTCTGACACACAAGGCATCGAGCAGCCTCTGAA
30 AGAACCAAAGCCACTAGCAGGCTTCCTGACTTGGGTTTGTAGGTAAGTAA

TACTCCCTTGAAAAATAAAAACATAGAGGCACTTTTCTCCTGGCTGTTTAT
TACAGAACGAAGAAAAAACACACTGGCTTGAAACAGACGCCAGATTTC
AATGTAGAGGTGAAATACGAGGTGGCAATTAAAATGTGATTACAGAAAG
TCTGGACACTGAGAAAAGTTTACAGGACAGTGGGTGTGGGTTTTCTATAA
5 CAGACACTTAAATATACATGACGATAATTGCAGATAGAAACCATCAAAGA
CAAACCCCAAATCAACTAATAATGTTTACAGATGTTCCCCCCCCAAACCAC
AGAGCCTTACATCAAAACAAATACTGAAAGGCTTTAAACCAGGAACAGCT
CGCCTTAACCCACGAGGGTGCACACAAGCTGGGCTTTTTCTCTCGGTCTG
AATGGTAAAGGGAGGAGGATACTCTAGCTCCTCCAGGTGGATTGCTGAGA
10 CAGGGCTCGGCTCACACACTGTCTCTGCGCCTCTCCCAAATCTGGAGAAC
TCTCCAGCCTCCTGGTAAAGTGTCTCTGTGGGGCACTTAACGATAAAAC
AGCTTCTGCTGTAAAGCTCATTAGGAAAGAGCTAGCGGAGACTGAAAGGT
TCGCAAAAGAGATTAAGAATCACACAAGGCAATAGGATTTTTAGTGAACA
TAGAAATAAATGGCCAAGTGGTTTTCTATTTGGCATTGTCAACTTGCACA
15 ACAACTCTTGGTCATATCCACATTGCTCATTGCATTAAAACCATAAGCGAC
TCAGCCACCTAGCTTAACAAGGTATCACTGGAGCAAACAACACGGTCTGC
ATATTTGTAACATTGTATAATAAACACAAAACAATGCATAGTAAACACAA
CTCTACTGAAACAAAAGCCGTCGCTTTATTTACAAAGTCACAAAATGAAG
TATAAATACTTCTGTCATTAATGTTTAGGAAAACCATTTACAAAATTTTCA
20 AATATGTACACGTAGCTTGAAAAATCACCAGCTTTCCATTTTGTACAGGT
AGAGAGAGGGATAAGCATGGGCTGACAACACCACTCAAATTGTAACGGG
AGACAACTGCGGGTATGGATCGACACCACTTCCTAGAGTGATGTCACCAT
GGGGGTTTCTATGGGCATCCTGCTCAGATTTAAAGTGCCCCAGCATCCTG
GGTGA CTTGCCCAGAATTCTGGGCTGTGGCATTTTGAGCAGCAGCATGCT
25 GTTCCAAAATGTCGTCGATCAGCCTCAAGTTGCACACCCAGTCTTCATCTG
GGCTCACACAGGAGCCTTTCAAGAGAGCTTCAATGAAATCTACCTCATTG
CAGTCAGGTGACGAAATCAGATCATTTAGTGGGGGTTGGGGCTGGCGCAA
AAAGTCGGCAGGTGGCAGCTCAGGGGGAATATCCGTTCTGTGCGAACGGAC
CTGGGAACTGGCTGGCAGCAACGGCAGAAGCAGCAGCAGCGGTGGCAGC
30 AGCAGCCACATAGCTTGGTGGCTCGATGCCCTGTATGGGGCTCAGGGGAC

TAAAGCTGGCCATACCCTGCTGGAGGAACTTGGTGGTGTGTTGCTACAGGC
ACCGGGCCCTGTACCGGGCTCTGCCTGAGGCTCTGGCTGCCCAGCAGGCT
GAAGCTGGGGTTGTTGGCCAGGGGCACTTGTGTTCCCATCGCAGCGGGCA
CTTGTGCCTCCCAATCAGATGGCCTCTGAAGGCAGGCCTGGCCAGAAGGT
5 GAGTGCTGCTGAACGCTATTATCCACTTGGCTGAGGGGTGTTTTCCCCGAA
ACTGCTGTGGTCACAGCTGCTGCCGCTGTGACCCATGCAGCATTGTTGAA
CGCAGTGGGCATTCTTGGCACACTAGGCCGTCTGAGCTGGTGGGGACTCA
AGGACTGGGTGCCCAGGGAGCTGGGACAGAACCCAGGCAGGGGCACTTC
TGGTGGGGTGGCCTTGGGGCTCTGCATATGCTGGCAGACAGAGTCAAGTC
10 TGCCCAGGGGAGTCTGGCCTGAGTGTGAGAGGATGGGACACTGGGGGCT
GGAGGTGAAAATTCCTTGCCGCTTCCCCAGAGTTGGTGAGATCACTCCCA
TGCCCGGCGCGCGCGTCGACGTGAGCAAGTGGAATGGAGTGGCAGGGAA
TGAGTGAGGGAGACAACAGCTCAGTAGGACCTTGTGGGCTACTGGAAG
ACTTTGGCTTTGACTCTCAGTGAGATAAGGAGGTATTTGAAGGAGTTTTAG
15 CAGAGAAGGAACAATTCTGTTTATTTTTAGGAGCTCTTTATATATTTAGGA
AATTACTTCTTTGTTAGTCATATATTTTGGAAATTTGTTGTTCCCCGTTTGTC
ATTTCTGTTTCTACATTTTGCTTATAATATTTGTCTATGTAAACCTTAATAT
GTACTTGGTTAAATTGATAAATTCCTTGCATAATTAATAGCTAGCTATTGT
TGGCCTACCTATAATTAATTTTTGCTTTATCTAGTAATAGGCCTGAATTTTC
20 ATTTGGGATCCAGCCCTCTCCCTGACACTATGCAGGGCCTGGGAGACTAA
AGCCTGCTGGAAATAACTCCAGCTTGGCCCTGCCAAACATTACAACACCT
AGCACCTGCTCTCCCCACCATAGTGACTGATTTGGGCAGGAATGAGACCA
CAGCCACTCCAAGAGCCCAGCTGTTGGCTACCAGTAACATCTTCTCCTCTA
AACAGGAGATCCAGCTATATTGAAGGCTTCCAAAATTACCCTATGTCTAC
25 CTGTCACAGGAATGTTTCAGAGGACTAAGCCTCTGCCTTTCTCCACCCCAT
CTTACTTCATATGGAACCTATGAATGGGCAACTTCACAATTTATTATTTACG
TCACCTAGATGCAAGTCAGAGGGAGAAGCAACAAATCAGCAGTACCTAA
GAAGTTCTCAAGTCCTCAAGGTGAACCTTACCTATATATACAAAATATATT
TGCTCTTTCTTTAACCCTTTTTCTTACTCCCACCACTTTCCACCCAATGT
30 GTAATAATATTAGTCACTTCATAGAGTTCTTATGAGAAGTTAAATGAGGA

AATCATGTAAACCACTGAGGGCTGGCAACAAGTTCTTAATAAGTGCTATT
GAGTTATGAGTATTATTATCTCCCTTGGGCTGGATTACAGTGTACCTCT
TCTGGGACCATAGTAAGCCAATGGTTCAGCTCCTTCTATTATGTCATTTT
AATTTGGGGAATAGCTTAACATATTAGATGATCCTTTACATGCCTGAGCG
5 GTCTTTCTTTGATGGTATTTCTATGATCTTATGAAGAGCTAGATGTTGCCG
CTCCATTTTATCTTGCTGGTGGCAGCTGGTGTITAGCCAACCTGTCTTTTTT
TTTCTGTCTTTATATTTTTGGTTTTGTTTTTTGGCAGAAGCCTGTGTTTTTCA
GATTACTTTTTGCTTGTGTCTACTTTTTCCCAAAACAATATTACATATTTAA
CTTATTGTATATTTAAGTTTATTTTATGCATCAAAAGTCGCATTGAATTTCT
10 GTGCTATTGTGGGGCAAATTGACATCAATGTAATTTAACTCATTTTATCCC
AAATTAGGCTATGTCTTTCCATATACTCAAATCTTCTTTTATGTCCCTCAAC
AGTTTCCTTCATATAAATATTGCATATTTCTAATTAAGTTTATTTTTCTAAT
CTCCATTTTGGATGTCTTAAATATAGTCTTTTCTCCCTTTGCTCTTTATCAA
AGATCTGCAAACCTATGGCCTGTTTTTGTAAATAAAGTTTTGTAATAAAGTT
15 TAAATGGCCACTCCCATTTGTTTACATGTTGAGTATGGCTGCTTTCAGACT
GAAATGGCGAATGAAGTGGTCATAACAAGACCATATGGCCTGGAAAGCA
AATATATTTATTATCTGGCAATTTACAAAAAAGTTTGCCAACCTCTGCTT
TATATTTTTGATCTAGTTGTTATTTGAAAATAAAGCTATTGATATTTGTATA
TTAATTTTTTACCCAGCTACCTTATTGTTTTTATTACTTAATAGTTTACAGT
20 ATTTCTTTTAGTTTTTCCAGGTGTAACATCTGTAAATATTAATAAATCTATT
AGCTAATATTTTATTTAGCATTTTATATCTATACTCACATGAGAAGTTTT
CTTTAATTTTCTTTTGTGTTTATCAACCTGTGTTCACTCAGGAAAATGAAG
CCACTTAAGTATACCAGGAGTGGAGGGTTTCATATAGAAATGAGGGCTTA
TTTGAATTTTGGGAAAGCTGATGAAACAGGTCAGCAGAGTTACTGCAGAA
25 GACCAGGAAAGTTACATCTGGAAGGTTAGGGAAGCAGACACTAAAGACT
TAAGCCTGAAGCATAGAAGTAAGGATCCGTCAATACTTATTGAGAACTT
CTGGAATTCTCAAGAACCTCTGAGAAATCTCCCATTTCTGTGGCAATAAAG
TGGTGGTACTCAAGCCATCACCAACATGGCTCATATGTCTCATTGCACTAT
GTCCAAAGACAAATGGCCTCTGCTTCTCTTCTGTCTCCTAAATCTCTTAAG
30 ATTCCTCTTATTTGTTTAAAGTCTAACCAGAAGACTTACATGGCAGGTAA

TTCTGGGAAATGTATTCCCTTGTTTCTCCAAAGGGGTGGGGATAATACCA
AGTTGACAACAGAAAATCTAGCACCCTTTTTAAGGTTTTGGCATCAGTGA
TTTCTTGGTGCTCTAAATAGTTCTTATGATTATTCCATGGGTGCTTAAAAA
GAATATGTATTCTCTGATTTAAGAATATGTAATTGATTAACTATTGATTA
5 GATATTAATAAAATAGATTTTTTAAATTACCTTATTAATTTGGTTAAGTCCT
CCATGACTACTTATTTTTTGACTATTTGTTATGGAAAGGTGTAGCTAAATA
AATATAGAACTACTCTATATGTGTTTCTAGTTCTCTATGTATTCTCTACATT
TTGCTTTATAATTGTTTTCTTTCTTAGAGACAGGGGTCTTGCTGTGTTGCCC
AGGCCAGAGTGCAGTGGCATGATCATAACTCACTGTAGCCTTAAACTCCT
10 GGGCTCAAGCAATCCTCCTGCCTCAGCTTCTCAAGTAGATGGGACTACAG
AAGTGCATCACCACACCCTGCTAATTTTTTTCTTATTTTTATTTTTAGTGAT
GGAGTCATGCTATGATGCTCAGAGTGGTCTTGAGCTCCTGGCCTCAAGCA
ATCCTCCCACCTCAGCCCCCTGAGTAGCTGGGATTATAGGTGAGAGGCAC
CATGCCTGGATTTGCTTTATAAATTTTGATATGTTACTTGGTGCATAGATA
15 TTCATTGCTGTTAGACCTTCATTATGTGTTGTACCCTGCCCTTCTTTGGTTT
ACTTTTCCTTGAATTCAGTCTTGTCTGATATTAATTTTCAGCTGTCTGCTTT
CTTCTTATTTCCATTTTTCTAATATGCTTTTGCCTACTCATTGCTTTCAATT
TTTTCTGAGTCATTTCACTTAGATGAGTATATTACAGTCAGTACTTAATT
GGGTTTTTCTTTGTCACTGAATATGAGTGTCTTTTTCTGTGTATTGATAG
20 GGTAGATATGTGTAGTATGAGTGTTACAATTTTTTAAAAAACTTTGATTTC
TCTCCTTTGTTTTATTTGGGGTCTCCGGTTATTTTTTAAGGTGTATAGTGTC
ATTTTATTTTTAATGTTTGCCTCTGTAAAAATAATTTTATAGGCTGTACTTA
ATCTTCTATTTATTTAGCTAGTTTATATTGACTCTTCAATAACAGTAGTAA
ATGTCCCTACTTCTAACTCCCCTTGCCCTCCACCTTTTAATTTTGTGGGGTA
25 TACTACTGGCCTTAGTTTTTCTAGTGGTTAATTTTATACATTTAAATAGGTT
TATACCTTTTAACTTCAATGGCCACCTTTAAAAATATTTTGACCCACCC
ACTATAAAAAATAAACATGCCATCTTACACGCACAATCTCACTTTCTCTCC
CCTTACATTCCTAATTTTTACTGGGTATTTAATATTTCTTCATCATCGTG
GTTGATTACATTTACATGTATATTGCAGTGGTAGTTCCCATAAATTGTTATA
30 ACCTTTAACCATAGTTCTACATTTATTTTTGCCACCGTTTTTCCCATTACC

TTTTGGTTCAACTTTATCTTCTAGTAGTGTCTTCAAAAAAGTGTCTTGGA
AACTACAGCCCATGAAACATTATTGCATATTTGAAAATATTTATCTGTCAC
CATTACACCTGAATGACAATTTAGCTAGCTCAAATTTACTGGCTCACATCA
TCTTTTTTTTTGGTAATTTTTTAAGTACTTCTCTACTGTCATGTAATGTTGTA
5 TGTAGGATAGCCTAAAGCCAACTTGTGTTTTTTTATTTTGCTGTTATAGTA
GTTTGATCATTTTTGCCTAGATGCCCCATAGGATTATTTTCTTTGTCTTTA
TTTCAGGAAACCTTACTAGACTATGGTTCAGTGCTGACTAGACATTTTTTT
CCTTAAGAAACACCATGCCCTTTAGATCCTTAGATTCAAATCTTTTTTAAT
ATCAGTAAAGTTTTTAAATTTACATATTTGGATATTTTATACTTCTCTAGTT
10 TTTCTTCTTAAAGAAGAAAACATTAATTCTGTCTTTGCCTGTCTTCCATATC
TACAATTTTCTTCTAATCACTTTTTTATAATTTCCATTTCAATTTTGCTCAATT
GCCTCAAAGCTGCTATCTATTTTTCTGTTGTCAACAATCTGTACTCTTACAT
CAGCTACTATGATAATATTGTGTCACCATTTACTACTATTTACAAGCTAAC
CATAATTTCTCTGATGATTTTAGTTTTCTCTTCTACTTTGTGTCCTAGATAC
15 AGGCAGCTCAATTTTAATTTTTGTGCTTTTCTTGATTCCCTTATACCTCTGC
ATTAAGCTCTTAATTCACAAATTACTTCCATTGTTAAGTTATTGAGTTCTT
AAACAAATATTTGGTCATGATTTTCATCTACTCCACATCAAAATGTTCATA
CTGTGTAGGCTGTTTGCCTTTTTTGTCTGCTATTTTCCACCATTTTGTGAC
ATTTTTGTATGGATCTCCTTGTTTCGTTTTTTATTATTAGTCATCCTTGAGTG
20 AGCAGAAATCTTCCTAGAGCAGGTATTTGAAAGAAGTAGTGTGGAAGAG
GGGCTGGGGCTGTGTTCTGCGTTAGTAGAAAATCCCCTAATTCTTAGAAA
AATCCTTTATGATTCATGGTCTTGTGGGAGAAGGTTGTGAAGCGGGTTTTG
TGAATGTGCTTCATGTGCAATAGAGAGTGACTGAGAGGTGGTTTAACCTT
TCTTCTTAGCTCACAAAGAATGGCTGCTACAGAGCTGCTTAACACCCACA
25 GTTTTTCTTCTACATGCTGCACATTTACAAAGATGGAGTATCCGTGAGTTT
TCTCACTCATTCCTACCTCCCCTGCCATTTTAAAGTATCCAAGGAAGTTGG
GGAGATTCCTCAGCTAAACAGCCACCCAGATCTTATTGTTCCAAACAAA
AAATTATTAGATTTGCCCTCTGAGCAGGTCACCTACTTTGGAGAATTTGT
GCTCTGCCAGCTCTCCTCAAGATTTGGAAAAGCAGTTATTTTCTATTGGAG
30 CCCCCATCACTCCAGTCCAAAGTTCAGTGCATGTGACATCCCATCTTATTT

GTCTGGTTCCTGATTATAGGAGTTTCCCAGTTTCAACAAAAATACTGTCTT
TGTTTTTGTCTCATTCATTTTTCCCCAATTTTGTTTCAAAGCGAAAAGTA
AAAACATTCTTACTCTAACATCTTAAAATTAATGTCTGGGAGGAACCTAA
ACTGCTCATTAATATATGCCATGCTATATGTGATATGGCAGAGACTGCTAC
5 CTGTCCACCAGAATCTATTTATTTTTCTTGGATGCACAATAAGACTACATT
TCCTAGTCTCCCTTGCAATTACATGTGACCATGTCATTAAGATTTAGTGGA
TGAAATTTAAGCAGAAGCAAGGTGTGTAAATTCTAGGGCTGGCCCATTAC
AACTTTAACATGGGCACCTGTGTGTTTTTCTTTTCTGCTTAATTGGGTG
GAGATAATCACAGTGATGCTGAAAGCCATGTGCTAAAGAAGGCAGAGCC
10 TCCTTCTGCTTGGGCTCCTGAATGACTGTGTGGAACAGAGGCCCTTGCCAC
TCTTGGGCCATTCTGGATTATTACATGCACAAAGAGTAAAGTTCTTTTCTG
TTAAGCCATTAGATGTATGGGTTTATTTGCTGTTACAGCCTAGTCGAACCC
AAATGATGTAGTGAGGGAGGCAGGACAAGGAAGAAACCTATTCAGTATG
GAGTTATCCTTAGGGACAAGTTACAAAAATTATAAGGATCAGACACCACT
15 GCTGGTTGATCCATTATAGACCAGAATAAGCAGGTGGTATGATTCCAGAA
GGAGATGCCATGGTTGTGTGAGAGAGAGATTAACCATTCCTATAAACAGA
AGGGTCAAAGAAAGTAGCCAGTCACCAGAATGCTTATATTAAGGAGGATT
GTGACTACCTGTATATAAAGTATAGGAAGAAAATGGGGCAGGGCAATTA
AGATGAATTGCAGACACAACCTATTTGTGCATATTAATTTGAATTTAAATTT
20 AAACCTAATTTAAGGATAATTTAAAGGAAAATCTTCCCTTACCTATTTTCA
GCTGAAGGGGTCTTTCAAGTTGTCAAACCTCTCTCTATATAAATAAGACTGC
ATAATCCAGACCCTCCGGAATCCAGAGATGAATTGCATGTGGTAGCTGAT
GTGTGGGAAGCCCACAGAGCAGCACTTCTGGTCATTGCCATCCAGCTAGC
TGGTGCTCTCAGAAGAATGAGATGAAGAGGCATGACTTTCAAGCTTCCTC
25 TGCCATACAAGGTGAGTTTGGGTTCCAGGCTTCTTCCTAGCCCCCTGTTTCT
GCTTCCTCTGATTTAATACCCAGTCACTGAGTACCTACTAAACACTCGGCA
TTAAACTGAAAAGAACAGGGAAGTCAAAAACCTTTTACCAATGCCTGATG
GGATAAAGACATGACCCATCCCTCAAAGAGCTTGTGGTCTTCCCATGGCA
ATGCCCTTCCTCCCTAGGCCAGATGGACTGTGAAGAACAGCAGTCTCAAG
30 TGGATTTATTGGGAAGCTCATGAAGACTCAGCTCAATGGGCCATGTTTGC

AGGGGCCCCCTTACAAGGCTCTGCACTTGATTTTGTAATTTTGAATTATTTT
TCTCAAAAAGGCCTCCCAAATTATGTAAGCTTCAGGCCTTTAAAACTTGG
ATCTGTTCCCTTGAAGCCACATAATCTGATCAGAAATAATTTTCAGGGCTC
ATTGGGTGCTGGAGAATTCATGGTTAGATATTTAAATAAAAAATGCAGAGA
5 ATCTATAGGACTTTCCAAGTGAAGACCAAGGCTTTTTTTTTCTGGTAGCTTG
ACTGCCAGGAAACAGAAAGCGGGGCATCTCCACCTGCTGGCAGAAGACC
AAAAGTACACCATATTGCATTGACCGCAGACATCAGGCTTTGTTTCCTTTT
AAGTGAATAGGTGAAATCAATCAAGCAAAGTGAATTAGGAAACAGAAG
GACCACTCACAACCGTACCTCCACAGCAAAGCTAACACTTTTCATTTTGAA
10 TAACCCCTTTCCAGCCCCTGTTTACATCAGTCACATTTTTTGCATGGACAT
AATAGTGTTCAGGCAGTCCTGCTCTACAGAAGGCATTTGTTTCGCTGACC
ATAAGCATCTTATGTGCTTCTACATTGTCTTTATAGCCAGCTGTCTGAAGT
TCTACATCATCCGTCAGTGATGGGTCCTGAGTAAACTCAGCAGTCTCCTCT
AAGTAACTTGTATATATCCCTTTTATAAAGAATGCTGCAGTGGCCACCCTT
15 AAGTATAGAACTTTGTCTTTCCCAGGCATTACTTACAGGAAACTGGGGAC
CAAATGCCACTCCATAGCATAGCTTAGGAGAAGTGAGATTCCAAAGGACA
GAGTGGGCTGGCACATATGTAGTAGAAGACTGGACAAATGAGAGTGAAT
GAGCGAGTCTGTTGTGAGTGGATGACACGCAATGCAACCCCAAGGGCCG
GCTCCAATTTGCACAGTCCCTAAAACATCCCAGGGATCCAGTCTTATCGC
20 AGTCTCATCAACATTGACAATAATCATTTAGAGAATTTTCACCATGAGACT
TTTGACACTCTGCTCTGGGATTTTGGAGTTCCAAGTCTTTAACAATTTAT
ACAGTCACTCAAATGTAGAATCCAGTGAATTATTAATGACAATTTTGCTCC
GCATATTGAAAGGGCATTTCTATATTAATCAATTATAAATCATGATAAATTG
CTTTCTAAAAATTATACACTTGAAACCCAGAAATGTTCAAGCCCCTCACC
25 AGTTTACATACATCCTTAACCGTGTCTTGTTAACTGGGTGACCAGGAAAC
AGAGAAAGGAGGATGTACTTAATGAATAACTGGCCAGAAGTCGGAAGAC
CTAGATTTTCTTCCCATCCCTTCCATAGATTTGCTCTGTGACCCTAGACAA
TTCATATGCCCCCTCTGTTACCATACTTTTAAGCTGAAGGGGTATAGTGC
ATACTTTCTGAGGTCCTTCTAACACTGGAGATCTATGAGTTCATAAGATCA
30 CATAGAGTCTTGGACTTTTCAGAGATCACTAGGCATAAGAAAGCCCTCGA

GGACATCCTCTTCCTAGATCAACCACTCCAAGCTAAGACCTTAATGGGTCT
ACACTGTCATCCATTTATATAGGAGAGAGGATGGTACAAGCTAAGTGAAA
GTCTGTTGTCAACTCAGGAACTAGGGTACAGTTCCAGCAATCAATCTCCA
TGGAGGGTTAGCAGAGGTCTGGCCTAGCCAATTAGGGTTTGGCACTGCCA
5 TATGCTACTAGAAAACACATGCTTTCATGCCAGGCAACAGGGTGGGAAGT
GCAATGCCATTAAGGACCATACTTAGCCTACTTAACTAGCCTGGGATT
GAGTTACTGAGCTCAAAATACAATTCAGCACACTTTCCTGCCCCTGCAA
GTCAGGCCAGCCCCTGGAGAACTCCAGGGACAAGGCAAGTCTGGAGTTTG
TCCCCAAGGAACTCCCAGTCTAGTGGAGGAGATGGGCCCATAGCAGAGTT
10 ACAAATGTCCAAAATAAGAGCACAAATTCCTGGGTCTTAAAGAGAGTTC
AAGCACACACACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAAAT
TAATTTAGATAAGGAAGGGAGGTAGGGAGAGATGGTGCCAGCAATGGCT
TCCTGAGGAAATGTCACCTTGTGATGGGTCTTTTAGGTGAGTTGGACATTGA
GAGAAAGTAATCAAAAGGGCACAAGATAGGGTAGCTGTCAGGTAAACAA
15 AACAAAACGTATGCACAAAGGTCTAGAGGCAAGAAAGCATGGGTGTGC
TTGCATAGCAATGAAGAATGTGGTTTGGCCAAAGCATAGGATTCACAGAA
GGACCTCCCCAGGCCATCAGTAGTCTCTCCTTCGCAGGAGTCCACTGTGG
CTGGACGAGGAACAGATCAAAACAGGTCTTGTTTGAGGCTTGGCCCAACA
ACAGGGCCAGCTGTTGACCAGGGCTTCCTTCACCCTTAGCTTTCTCTCCCC
20 CTAGTGTAGTCAGAGACTCCAAAGGCAGTCCCAGAGTGTTGTGCCAAAAC
AAATGCACATCAGGCTATTGTCATCAGGAGCCCTTGTCTTGTTCTGTTTGT
TTTAAGAAGACATAAGTGAAATAACCAGGTTTCCGAAGAGAGGGCAGAA
TTGGGTTTGTCCCCAGGGAAAATGAGGTTTCTCCACCATGGCTGTTCTCCC
CATGCAGCAGCGGCCAGGGGTCCCCAGCAGCTCTTTGCTCAGTGCCCCTG
25 ATGCTAGGCCTGATGCTTCCTATGTGGTCAGCTCAGATGTGCAACCCTTCT
CCATTTTTATTATTTGAGCAATATAATCATTATTGCTTTTGTACCCAAAAT
CTTATAATCACCCATTTAAGAGTAATGTAGGAATCATTTTTCTCTTTTATTT
GATGAACTAATCTTTCTGGTTGGCTCTTGAGGTGGGAGGGGTGGGCAATG
GGGACCAGCCTGCTTCATCTCACTCTTTCCTAGGCCTGCTCCTTAGGCACG
30 TGCCCCACAAATCCCACCTCCGCAGGTTGGAAAAGGGGCCACAGAGGAT

CTCAGGTATCTGAGGAAGGGCCAGAGGCTATTTGGCCTGACATCCTGGAA
CATCACCAGCCAGGAGCCATAACTCCTTTAGGCTAGCCTGCCTGCAACTG
GATCAGGTATAGCAAAAACGATTGATCTCACCTTCTAATTTTCAATGAGT
GGGAGTGGCTGCCTTGAGTGCTGTGTTGGGAAGGAGTCTGAGGCAGCATC
5 TGGACCCTGTGTAAATATGCGTTTCCTGTTATTGATGAAGCTTTCAAGACA
AGATTTTAGAAAAGCTTTGGCTCAAGTTGTAGGTTGTGGGAGGAAGGAATT
AGTAAGAGCCATCATAAACTCAACATCTATTGTGTTCCAGGAGCTGTGAC
CACTCCCTTCTTTTAGTCGGGAAGGTCAGGATCTCACTATGTCACCTAGG
CTAGAGTGCAATGGTGTGATTATTGCTCATTGTAGTCTTGAATTTGAATTC
10 CTGGGCTCAAGTAATCCTCCTGCCTCAGCCTTCTGAGTAGGTGGGACTACT
GGTGTGTGCCACCACACTCGGCTTACTCCTCCCTTTGTATCCTGCTGCATC
GTTAGTGCATGATACCCACCACCCCTCAGAGGCCAGCCTAGTGCTGTC
TCTCAGAAAAACACCCTTTGAGTCAACAGAGACAACTTAGGGGGTGCCT
GACTCCAAAGTCCAGGCCTTCCTCCTTTGAGGAACAGCATGGCTTCCCCA
15 GGGAATATCATTCCATGGAGACAAGGCCTGCAGACAGATCTCATTTTCAT
CTGTTCTTTTACTAGCCCTTGGAGGATTTGCTCCAAAAACCTGAGGCACTG
ATGTTTACTCCAGCTTTACTTCCCCCATACCTGGTGAACAACTCTGGCCCA
GCTTCCACACACCAAGGCCCATGGCTCTAATAGAGCCAGAGAGGCAGTCT
CAGCTAAGAAGAGCTGTATCCTCCAGGGACCAGCAGGAACCTTCCAGAGAT
20 ATCCTACAAGTCCAACCTGCTGGGAAAAGCAGACTGGGGTCTTCCTCAGG
CTCCACCAAAGATAGCCAAGTGGCAGCAATGACATGAGCACCTCATTCTT
GTCCTAGACCCTGGCAATTCCATGAGAGGGAGGCTCTGAATTCACCTCTG
TTCCCAGGATGTGTGACATCAATGCACATTGTCATACCAGATACACCTGC
CAAACCTTCAGAACCCCTGGGAAGTGACCAGAATGCCACTGAATTCAACAC
25 CTTAGACCAATGAACATGACCTGGGTCCTGTAGGATAGGAGGCAGCAGCA
GTGTATACCCAGACCTGGTCGACTCGAATAGATCTGCTAGGCGCCAAGCT
TGTCAGCCTTACCAGTAAAAAAGAAAACCTATTAAAAAAACACCACTCGA
CACGGCACCAGCTCAATCAGTCACAGTGTA AAAAAGGGCCAAGTGCAGA
GCGAGTATATATAGGACTAAAAAATGACGTAACGGTTAAAGTCCACAAA
30 AAACACCCAGAAAACCGCACGCGAACCTACGCCCAGAAACGAAAGCCAA

AAAACCCACAACCTTCCTCAAATCGTCACTTCCGTTTTCCACGTTACGTAA
CTTCCCATTTTAAGAAAACACTACAATTCCCAACACATACAAGTTACTCCGCC
CTAAAACCTACGTCACCCGCCCCGTTCCACGCCCCGCGCCACGTCACAA
ACTCCACCCCCTCATTATCATATTGGCTTCAATCCAAAATAAGGTATATTA
5 TTGATGATGTTT

Other embodiments are within the following claims. While several embodiments have been shown and described, various modifications may be made without departing from the spirit and scope of the present invention.

WHAT IS CLAIMED:

1. A nucleic acid molecule comprising a low homology packaging signal cassette flanked by a recombinase recognition sequence, wherein said
5 packaging signal cassette comprises a modified adenovirus packaging signal, provided that said modified packaging signal has low homology to a wild-type adenovirus packaging signal.
2. The nucleic acid of claim 1, wherein said recombinase recognition
10 sequence is *loxP*.
3. The nucleic acid of claim 1, wherein said recombinase recognition sequence is *frt*.
4. The nucleic acid of any one of claims 1-3, wherein said modified
15 packaging signal is less efficient than said wild-type packaging signal.
5. The nucleic acid of claim 4, wherein said wild-type packaging signal is human adenovirus serotype 5 packaging signal.
20
6. The nucleic acid of claims 5, wherein the modified packaging signal comprises at a maximum, 23 bp of contiguous sequence homology with said wild-type packaging signal.
7. The nucleic acid of claim 5, wherein said modified packaging
25 signal is about 2-3 times less efficient than said wild-type signal.
8. The nucleic acid of claim 6, wherein said modified packaging signal comprises two to six A elements, each A element having a consensus sequence
30 of ATTTGN₈GC.
9. The nucleic acid of claim 6, wherein said nucleic acid is a plasmid.

5

10

15

15. The nucleic acid of claim 14, wherein said insertion is a human intron sequence.

20

25

(a) an adenovirus genome having an E1 region deletion;

(b) an excisable packaging signal cassette replacing a wild-type packaging signal, the excisable packaging signal cassette comprising a 5' *loxP* site, a modified packaging signal and a 3' *loxP* site, wherein said modified packaging signal has low homology to and is less efficient than the wild-type packaging signal; and

(c) an optional insertion element comprising at least about 2900 base pairs of non-adenoviral DNA inserted in the E3 region without deleting any part of the E3 region.

30

17 . The virus of claim 16 wherein said adenovirus genome is human Adenovirus type 5.

18. The virus of claim 17 wherein said modified packaging signal comprises at a maximum, 23 bp of contiguous sequence homology with said wild-type packaging signal.

5 19. The virus of claim 18, wherein said modified packaging signal is about 2-3 times less efficient than said wild-type signal.

20. The virus of claim 19, wherein said modified packaging signal comprises two to six A elements, each A element having a consensus sequence of
10 ATTGNGC.

21. A vector comprising the adenovirus of claim 16.

22. A cell line expressing E1 and infected with the helper virus of
15 claim 16.

23. The cell line of claim 22, wherein said cell line further expresses *cre* recombinase.

20 24. The cell line of claim 23, wherein said cell line is 293 *cre* cells.

25. A helper-dependent adenovirus vector comprising:

a) a 5' ITR;

b) a packaging signal;

25 c) at least one heterologous expression cassette;

d) a human genomic stuffer DNA;

e) an optional E4 non-coding segment conferring a selective advantage, wherein the E4 element is located between nucleotide -400 from the right end; and

30 f) a 3' ITR;

wherein the overall size of the vector is between about 28 kb and 36 kb, and wherein the only adenoviral sequences present are the ITRs, optional E4 non-coding segment and the packaging signal, and wherein no bacterial origin of replication or bacterial marker genes are present.

26. The helper dependent adenovirus vector of claim 25, wherein said optional E4 element is present.

5 27. A plasmid vector comprising
a) a 5' ITR;
b) a packaging signal;
c) at least one heterologous expression cassette;
d) a human genomic stuffer DNA;
10 e) an optional E4 non-coding segment conferring a selective
advantage, wherein the E4 element is located between nucleotide -400 from the right
end; and
f) a 3' ITR.

15 28. A helper-dependent adenovirus comprising in a 5' to 3' direction:
a) a 5' ITR,
b) a packaging signal cassette directly joined to the 3' of said 5' ITR,
c) a first stuffer DNA at least about 1 kb,
d) at least one heterologous expression cassette,
20 e) a second stuffer DNA at least about 1 kb,
f) an optionally present non-coding E4 segment at least 400 bp in
length; and
g) a 3' ITR, wherein said 3' ITR is directly joined to the 5' end of said
non-coding E4 segment if present;

25 provided that said helper-dependent adenoviral vector does not encode
one or more proteins needed for viral generation, is about 28 kb to about 36 kb, and
has a GC content between about 50% and about 60%.

29. The helper-dependent adenoviral vector of claim 28, wherein said
30 virus is between 30 and 36 kb in length.

30. The helper-dependent adenoviral vector of claim 29, wherein said
first stuffer and said second stuffer are derived from inverted mammalian non-gene or
intron sequences.

31. The helper-dependent adenoviral vector of claim 30, wherein said virus does not encode for any adenovirus proteins.

5 32. The helper-dependent adenoviral vector of any one of claims 28-31, wherein said optionally present non-coding E4 region is present.

10 33. The helper dependent virus of claim 32, wherein said GC content is between 52% to 57%.

34. A method of generating helper-dependent adenoviral gene vectors in a cell line expressing E1 and *cre* recombinase comprising:

15 a) infecting said cell line with a helper-dependent vector comprising: a 5' ITR, a packaging signal, at least one heterologous expression cassette, human genomic stuffer DNA and a 3' ITR, wherein the overall size of the helper-dependent vector is between about 28 kb and 36 kb, and wherein no functional adenoviral coding sequences and no bacterial origin of replication or bacterial marker genes are present;

20 b) infecting the cell line with a helper virus comprising: an adenovirus genome having an E1 region deletion; an excisable packaging signal cassette replacing a wild-type packaging signal, the excisable packaging signal cassette comprising a 5' *loxP* site, a modified packaging signal and a 3' *loxP* site, wherein the modified packaging signal has low homology to and is less efficient than the wild-type packaging signal; and an optional insertion element comprising at least about 25 2900 base pairs of non-adenoviral DNA inserted in the E3 region without deleting any part of the E3 region; and

c) obtaining the generated helper-dependent viral vectors.

30 35. A method of generating a helper-dependent adenoviral vector comprising:

a) producing a cell comprising

(i) *trans* functions needed for adenovirus generation; and

(ii) said helper-dependent adenoviral vector, wherein said helper-dependent adenoviral vector comprises the necessary *cis* functions needed for

adenovirus generation and at least one heterologous expression cassette, and said helper-dependent adenoviral vector does not encode for any adenovirus proteins, is about 28 kb to about 36 kb, and has a GC content between about 50% and about 60%, and

5 b) generating said helper-dependent adenoviral vector.

36. The method of claim 35, wherein late proteins and either E2 proteins or E4 proteins, or both E2 proteins and E4 proteins are supplied by a helper virus present in said cell.

10

37. The method of claim 36, wherein said helper virus comprises a low homology packaging signal cassette flanked by a recombinase recognition sequence, wherein said packaging cassette signal comprises a modified adenovirus packaging signal having low homology to a wild-type adenovirus packaging signal.

15

38. The method of claim 37, wherein said recombinase recognition sequence is *loxP* and said cell expresses Cre recombinase.

39. The nucleic acid of claim 37, wherein said recombinase recognition sequence is *frt* and said cell expresses Flp recombinase.

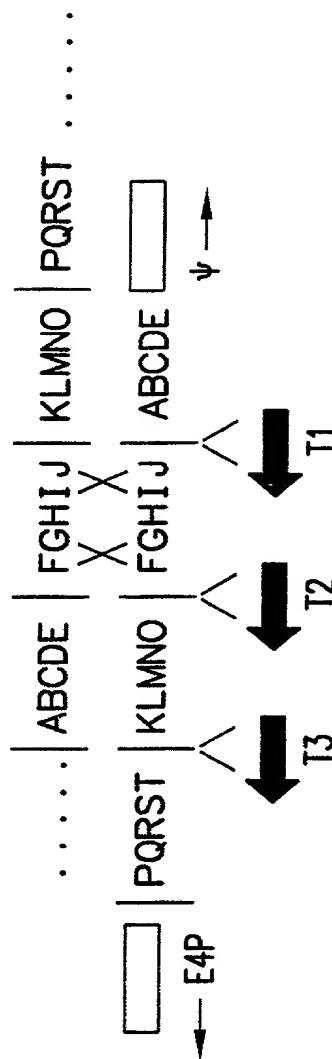
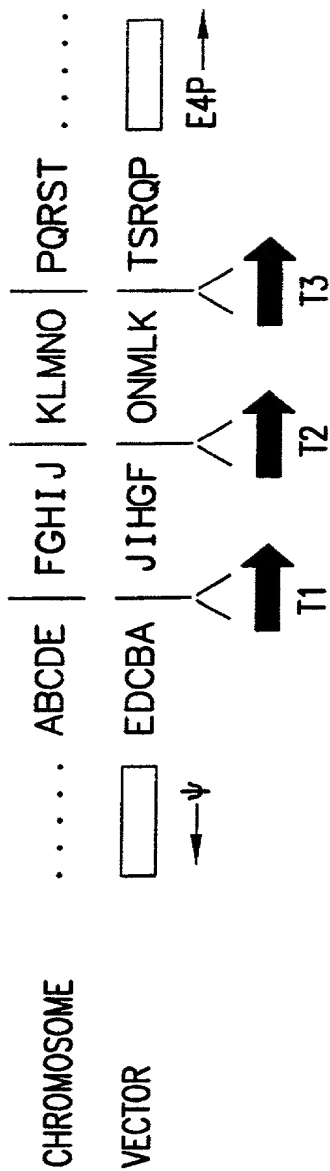
20

40. The method of claim 37, wherein said modified packaging signal is less efficient than said wild-type signal.

41. The method of any one of claims 36-40, wherein said wild-type packaging signal is from human adenovirus serotype 5.

25

PREVENTION OF INTEGRATION (BY HOMOLOGOUS RECOMBINATION)



- **CROSS OVER WILL INVOLVE EXCHANGE OF HOMOLOGOUS DNA WITHOUT INSERTION OF TRANSGENES.**

FIG. 1

3/8

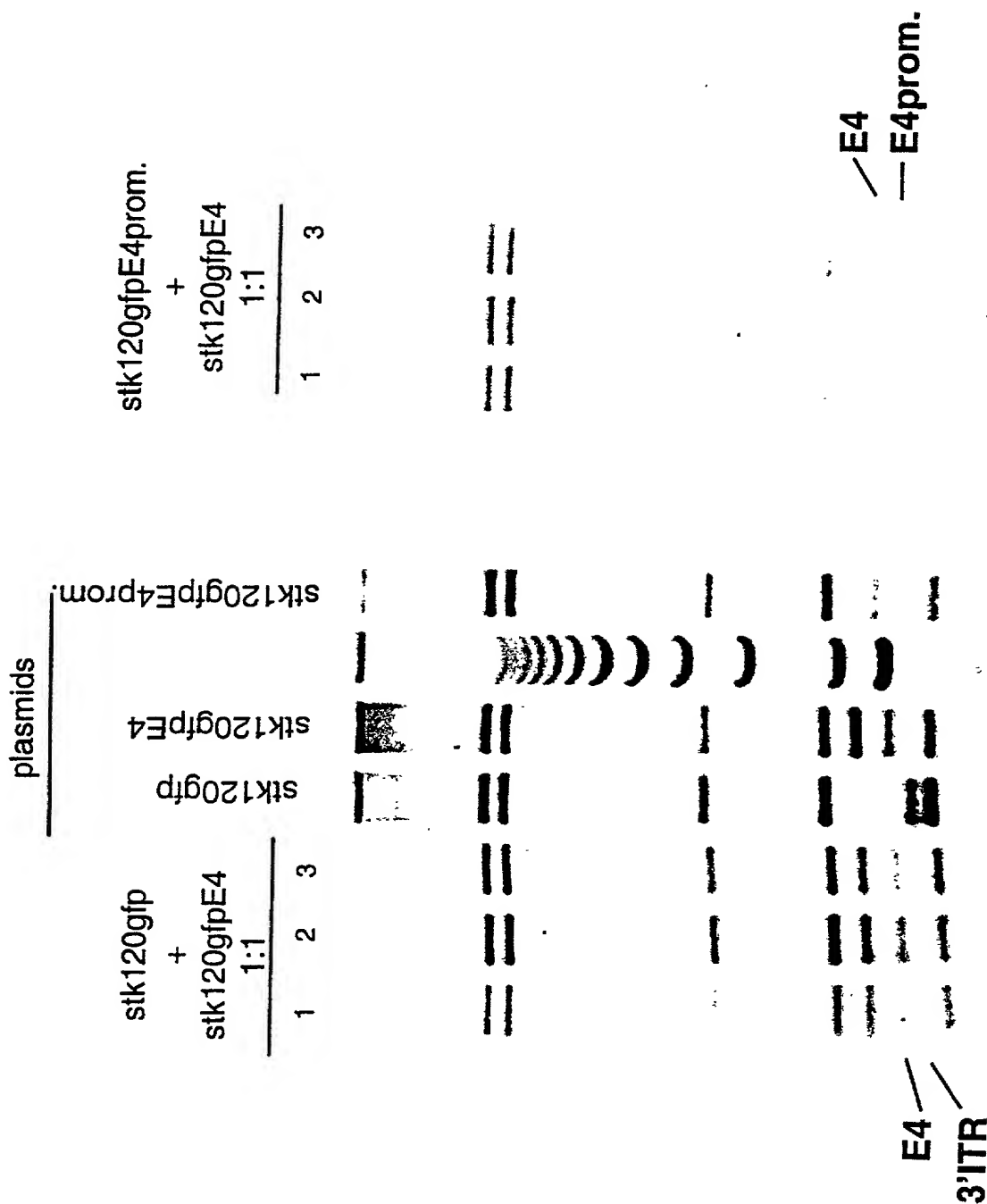


FIG. 2B

FIG. 2C

4/8

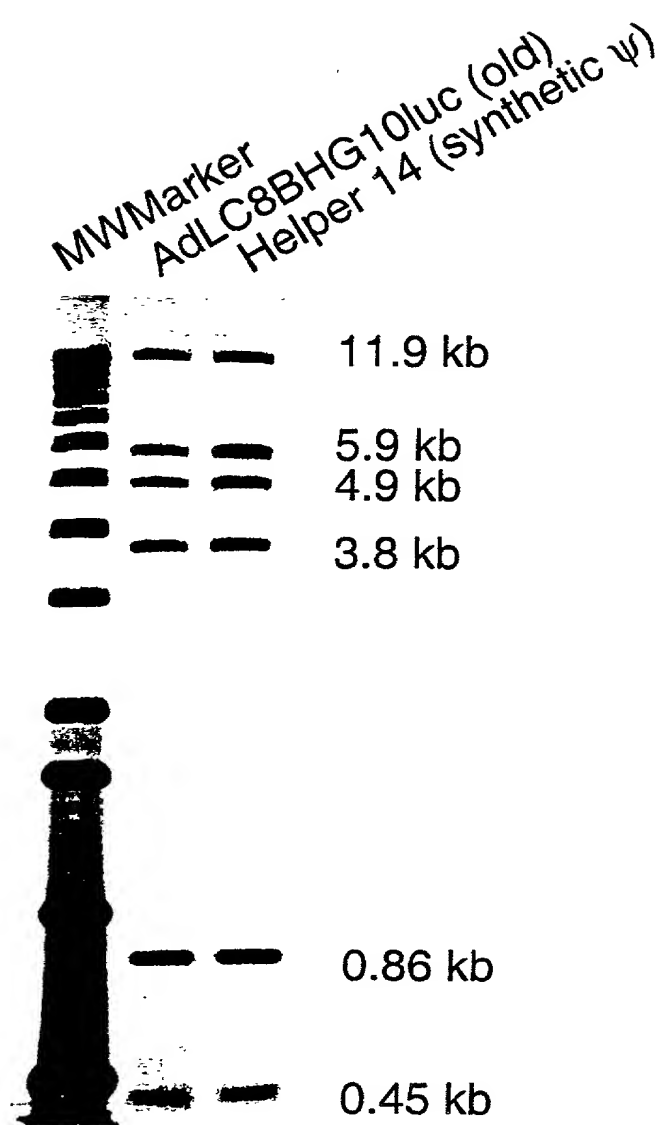


FIG.3

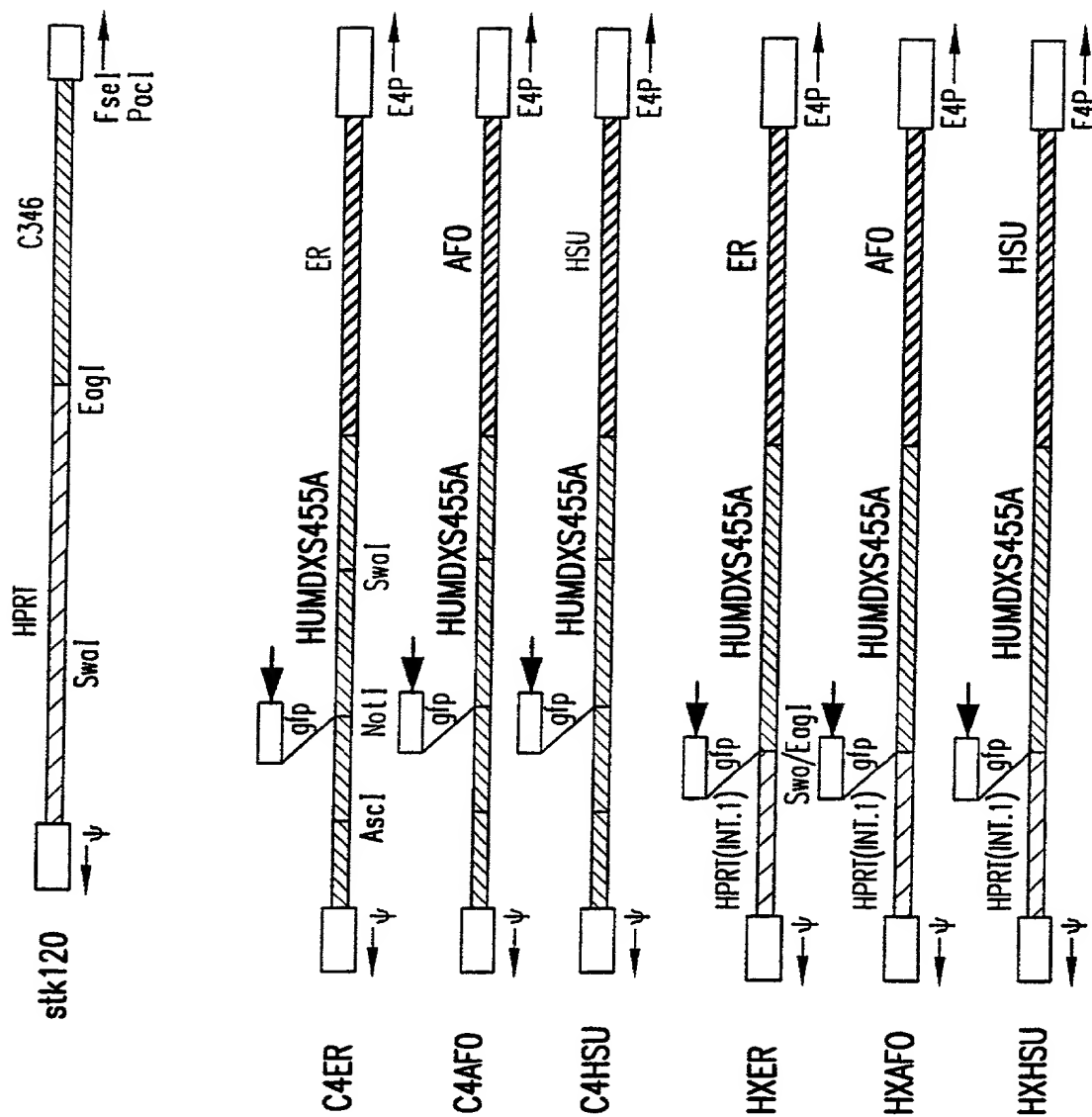


FIG. 4

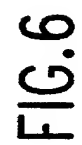
C4 HX Mix
HSU HSU

HX HX HX
ER HSUAFO
Mix

C4 C4 C4 Mix
ER HSU AF0

FIG. 5B

FIG. 5A



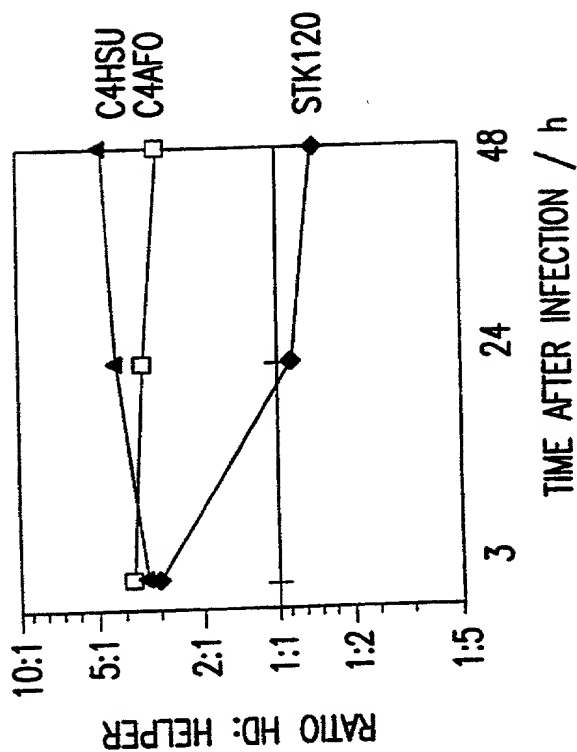
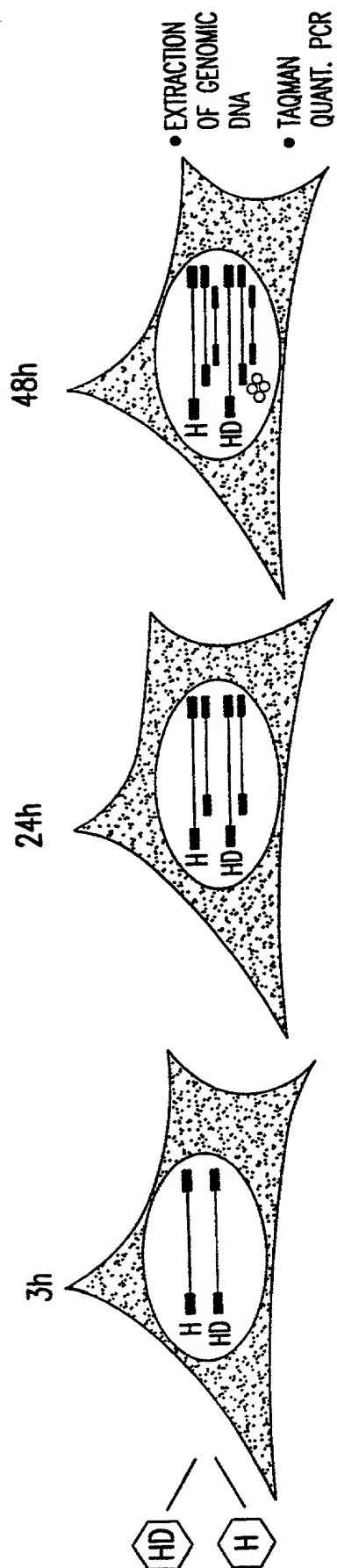


FIG. 7

**DECLARATION AND
POWER OF ATTORNEY
FOR UTILITY OR DESIGN
PATENT APPLICATION
(37 CFR 1.63)**Declaration
Submitted
with Initial
Filing

OR

Declaration
Submitted after Initial
Filing (surcharge
(37 CFR 1.16 (e))
required)

Attorney Docket Number

20377YP

First Named Inventor

A. BETT, ET AL.

COMPLETE IF KNOWN

Application Number

TO BE ASSIGNED

Filing Date

HEREWITH

Group Art Unit

TO BE ASSIGNED

Examiner Name

TO BE ASSIGNED

As a below named inventor, I hereby declare that:

My residence, post office address, and citizenship are as stated below next to my name.

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled:

IMPROVED HELPER DEPENDENT VECTOR SYSTEM FOR GENE THERAPY

(Title of the Invention)

the specification of which



is attached hereto

OR

was filed on (MM/DD/YYYY) as United States Application Number or PCT InternationalApplication Number and was amended on (MM/DD/YYYY) (if applicable).

I hereby state that I have reviewed and understand the contents of the above identified specification, including the claims, as amended by any amendment specifically referred to above.

I acknowledge the duty to disclose to the Patent and Trademark Office all information known to me to be material to patentability as defined in 37 CFR 1.56.

I hereby claim foreign priority benefits under 35 U.S.C. 119(a)-(d) or 365(b) of any foreign application(s) for patent or inventor's certificate, or 365(a) of any PCT international application which designated at least one country other than the United States of America, listed below and have also identified below, by checking the box, any foreign application for patent or inventor's certificate, or of any PCT international application having a filing date before that of the application on which priority is claimed.

Prior Foreign Application Number(s)	Country	Foreign Filing Date (MM/DD/YYYY)	Attorney Docket Number	Priority Claimed?	
				YES	NO
				<input type="checkbox"/>	<input type="checkbox"/>
				<input type="checkbox"/>	<input type="checkbox"/>
				<input type="checkbox"/>	<input type="checkbox"/>
				<input type="checkbox"/>	<input type="checkbox"/>

☐ Additional foreign application numbers are listed on a supplemental priority data sheet PTO/SB/02B attached hereto:

I hereby claim the benefit under 35 U.S.C. 119(e) of any United States provisional application(s) listed below.

Application Number(s)	Filing Date (MM/DD/YYYY)	Attorney Docket Number
60/118,601	02/04/1999	20377PV
60/138,134	06/08/1999	20377PV2

DECLARATION AND POWER OF ATTORNEY for Utility or Design Patent Application

I hereby claim the benefit under 35 U.S.C. 120 of any United States application(s), or 365(c) of any PCT international application designating the United States of America, listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States or PCT international application in the manner provided by the first paragraph of 35 U.S.C. 112, I acknowledge the duty to disclose information known to me to be material to patentability as defined in 37 CFR 1.56 which became available between the filing date of the prior application and the national or PCT international filing date of this application.

U.S. Parent Application or PCT Parent Application Number	Parent Filing Date (MM/DD/YYYY)	Parent Patent Number (if applicable)

☐ Additional U.S. or PCT international application numbers are listed on a supplemental priority data sheet PTO/SB/02B attached hereto.

As a named inventor, I hereby appoint, respectively and individually, as my attorneys or agents with full power of substitution and revocation, the following registered practitioner(s) to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith:

☐ Customer Number

OR

☒ Registered practitioner(s) name/registration number listed below

Place Customer Number
Bar Code Label here

Name	Registration Number	Name	Registration Number
JOANNE M. GIESSER	32,838	JACK L. TRIBBLE	32,633

Direct all correspondence to: ☒ Customer Number or Bar Code Label

000210

Name	JOANNE M. GIESSER				
Address	Merck & Co., Inc. - Patent Department				
Address	P.O. Box 2000, RY60-30				
City	Rahway	State	NJ	ZIP	07065-0907
Country	USA	Telephone	(732)594-3046	Fax	(732)594-4720

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under 18 U.S.C. 1001 and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Name of Sole or First Inventor:

☐ A petition has been filed for this unsigned inventor

Given Name (first and middle [if any])

Family Name or Surname

ANDREW

BETT

Inventor's
Signature

Date

07/18/2001

Residence:
City

LANSDALE

State PA

Country US

Citizenship

CANADA

Post Office
Address

Merck & Co., Inc., P.O. Box 2000

City

Rahway

State NJ

ZIP

07065-0907

☒ Additional inventors are being named on the 1 supplemental Additional Inventors(s) sheet(s) PTO/SB/02A attached hereto.

DECLARATION AND POWER OF ATTORNEY

ADDITIONAL INVENTOR(S)
Supplemental Sheet

Name of Additional Joint Inventor, if any:		<input type="checkbox"/> A petition has been filed for this unsigned inventor					
Given Name (first and middle [if any])				Family Name or Surname			
VOLKER				SANDIG			
Inventor's Signature	x <i>Volker Sandig</i>					Date	x 07/11/01
Residence: City	BERLIN	State		Country	GERMANY	Citizenship	GERMAN
Post Office Address	Merck & Co., Inc., P.O. Box 2000						
City	Rahway	State	NJ	ZIP	07065-0907		
Name of Additional Joint Inventor, if any:		<input type="checkbox"/> A petition has been filed for this unsigned inventor					
Given Name (first and middle [if any])				Family Name or Surname			
RIMA				YOUIL			
Inventor's Signature	x <i>R. Youil</i>					Date	x 07/19/01
Residence: City	NORTH WALES	State	PA	Country	US	Citizenship	AUSTRALIA
Post Office Address	Merck & Co., Inc., P.O. Box 2000						
City	Rahway	State	NJ	ZIP	07065-0907		
Name of Additional Joint Inventor, if any:		<input type="checkbox"/> A petition has been filed for this unsigned inventor					
Given Name (first and middle [if any])				Family Name or Surname			
Inventor's Signature						Date	
Residence: City		State		Country		Citizenship	
Post Office Address	Merck & Co., Inc., P.O. Box 2000						
City	Rahway	State	NJ	ZIP	07065-0907		
Name of Additional Joint Inventor, if any:		<input type="checkbox"/> A petition has been filed for this unsigned inventor					
Given Name (first and middle [if any])				Family Name or Surname			
Inventor's Signature						Date	
Residence: City		State		Country		Citizenship	
Post Office Address	Merck & Co., Inc., P.O. Box 2000						
City	Rahway	State	NJ	ZIP	07065-0907		